

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 12:17:06 ; Search time 350 Seconds
(without alignments)
8975.822 Million cell updates/sec

Title: US-09-828-592-6

Perfect score: 1395

Sequence: 1 atgattccatgtagtag.....ccaacccttgcttaagtaa 1395

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.GeneSeq_101002:*
1: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA1980.DAT:*
2: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA1981.DAT:*
3: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA1982.DAT:*
4: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA1983.DAT:*
5: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA1984.DAT:*
6: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA1985.DAT:*
7: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA1986.DAT:*
8: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA1987.DAT:*
9: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA1988.DAT:*
10: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA1989.DAT:*
11: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA1990.DAT:*
12: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA1991.DAT:*
13: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA1992.DAT:*
14: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA1993.DAT:*
15: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA1994.DAT:*
16: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA1995.DAT:*
17: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA1996.DAT:*
18: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA1997.DAT:*
19: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA1998.DAT:*
20: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA1999.DAT:*
21: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA2000.DAT:*
22: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA2001A.DAT:*
23: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA2001B.DAT:*
24: /SID82/gcgdata/geneSeq/geneSeq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1395	100.0	1395	24	ABK13828
2	1393.4	99.9	1489	12	AAQ10323
3	1391.8	99.8	1525	21	AAZ29808
4	1391.8	99.8	1599	22	AAFS4046
5	1390.2	99.7	1599	4	AAQ30203
6	1388.6	99.5	1395	14	AAQ50415
7	1387	99.4	1444	22	AAQ83246
8	1375.6	98.6	1467	20	AAZ32160
9	1295.6	92.9	1382	19	AAV41731

10	1294.2	92.8	1299	19	AAV41727
11	1191.8	85.4	1123	23	AA83474
12	512.4	36.7	1724	23	AA83475
13	397.4	28.5	578	22	ABA63081
14	390.2	28.0	395	22	ABA75533
15	340	24.4	433	23	AA83471
16	323.8	23.2	368	23	AA83473
17	267.4	19.2	326	23	AA83472
18	254	18.2	355	24	ABN5082
19	254	18.2	358	24	ABN5080
20	179.8	12.9	238	20	AAZ2186
21	139.8	10.0	1492	19	AAV8747
22	139.8	10.0	1492	19	AAV8748
23	132	9.5	1414	19	AAV8745
24	132	9.5	1414	19	AAV8746
25	131	9.4	1393	17	AAV35221
26	131	9.4	1393	21	AAZ39750
27	131	9.4	1626	24	ABK8415
28	128.8	9.2	1358	19	AAV8741
29	128.8	9.2	1358	19	AAV8742
30	126.6	9.1	1454	19	AAV8749
31	126.6	9.1	1454	19	AAV8750
32	126.4	9.1	1584	19	AAV8740
33	126.4	9.1	1584	19	AAV8739
34	126.4	9.1	2944	20	AA87831
35	126.4	9.1	2944	21	AAZ49271
36	126.2	9.0	1838	19	AAV8743
37	126.2	9.0	1838	19	AAV8744
38	124.6	8.9	1325	24	ABK83894
39	124.6	8.9	1425	21	AAZ35220
40	124.6	8.9	1425	21	AAZ39749
41	119	8.5	1316	16	AAQ75377
42	119	8.5	1316	17	AAV14255
43	119	8.5	1316	24	ABK84121
44	119	8.5	1316	24	ABL69984
45	119	8.5	1931	21	AACT7964

ALIGNMENTS

RESULT 1	ABK13828
ID	ABK13828 standard; DNA; 1395 BP.
AC	ABK13828;
DT	08-MAY-2002 (first entry)
DE	DNA encoding human antithrombin III (AT3).
KW	Human; antithrombin III; AT3; H-helix; AT-pos; coagulation deficiency;
KW	haemophililia; factor VIII; thrombin; T-TW; anticoagulant; hemostatic;
KW	gene; ds.
XX	
OS	Homo sapiens.
XX	
FN	US2001055591-A1.
XX	
PD	27-DEC-2001.
XX	
PF	06-APR-2001; 2001US-0628592.
XX	
FR	07-APR-2000; 2000US-195872P.
XX	
PA	(WALS/) WALSTON T.
PA	(COOP/) COOPER S.
PA	(REZA/) REZAE A.
XX	
PI	Walston T, Cooper S, Rezaie A;
XX	
DR	WPI; 2002-170988/22.
XX	

DNA encoding the m
DNA encoding novel
DNA encoding novel
Human foetal liver
Human foetal liver
DNA encoding novel
DNA encoding novel
DNA encoding novel
Gene #1580 used to
Human gene signatu
Human antithrombin
CDNA nESP15-1492 e
Complementary str
CDNA nESP14-1414 e
Complementary str
Cytoplasmic antip
Human cytoplasmic
Human CDNA differe
CDNA nESP12-1358 e
Complementary str
CDNA nESP16-1454 e
Complementary str
Complementary str
CDNA nESP11-1584 e
Mouse neuroserpin
Murine neuroserpin
CDNA nESP13-1838 e
Complementary str
Human CDNA differe
Human cytoplasmic antip
Human cytoplasmic
Human elastase inh
Human elastase inh
Human CDNA differe
Pancreas cancer re
Human cancer assoc


```

FT      /*tag= c
XX      /note= "Bock et al, Nucl.Acids.Res 10, 8113, 1982"
XX      W09100291-A.
XX      10-JAN-1991.
XX      23-JUN-1990; 90MO-EP01026.
XX      26-JUN-1989; 89EP-0201675.
XX      (ALKU) AKZO NV.
XX      DiJkema R, Visser A;
XX      WPI; 1991-036710/05.
XX      New modified antithrombin III variants - with altered
XX      heparin-dependent effect to factors IIA and Xa
XX      Example; Fig 1; 24pp; English.
XX      A human fetal liver cDNA library was screened with a conserved
XX      serpin probe and partial antithrombin III clone AT11cDNA1 was
XX      isolated. This insert was then used to screen an human adult liver
XX      cell library. A second partial clone was isolated (AT11cDNA2). The
XX      incomplete leader sequence was replaced by a synthetic DNA leader
XX      based on the published AT11 sequence. An EcoRI restriction site was
XX      positioned upstream of the start at -10 (= position 1 of this
XX      sequence). Mutations were introduced into the P-region, encoding
XX      amino acids 384 to 396, to give modified antithrombin III proteins.
XX      See also AAR10380-R10397.
XX      Sequence 1489 BP; 394 A; 372 C; 358 G; 365 T; 0 other;
XX      Query Match 99.9%; Score 1393.4; DB 12; Length 1489;
XX      Best Local Similarity 99.9%; Pred. No. 0;
XX      Matches 1394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY      1 ATGTATTCATATGATGAGAACTGTAACCTCTGGAAAAAGAGTTTATCTTTTGTCC 60
DB      11 ATGTATTCATATGATGAGAACTGTAACCTCTGGAAAAAGAGTTTATCTTTTGTCC 60
QY      61 TTGCTGCTCATGTCCTTCTGGGACTGCTGACCTGTCAAGGAGCCCTGTGACATCTGC 70
DB      71 TTGCTGCTCATGTCCTTCTGGGACTGCTGACCTGTCAAGGAGCCCTGTGACATCTGC 70
QY      121 ACAGCCAGCCCGGGGACATTCCTCATGATCCCATGTCATTAACGCTCCCGAGAG 180
DB      131 ACAGCCAGCCCGGGGACATTCCTCATGATCCCATGTCATTAACGCTCCCGAGAG 180
QY      181 AAGCCACTGAGATGAGGCTCAGAAAGAAATCCCGAGAGCCCAACCGGCTGTTC 240
DB      191 AAGCCACTGAGATGAGGCTCAGAAAGAAATCCCGAGAGCCCAACCGGCTGTTC 240
QY      241 TGGGAATGTCGCAAGGCAATTCCTGCTTGTCTACCACTTTATACACACTTGGAGAT 300
DB      251 TGGGAATGTCGCAAGGCAATTCCTGCTTGTCTACCACTTTATACACACTTGGAGAT 300
QY      301 TCCAGAAATGACATGATTAACATTTTCTGTCACCCCTGAGATCTCCACGGCTTTTGT 360
DB      311 TCCAGAAATGACATGATTAACATTTTCTGTCACCCCTGAGATCTCCACGGCTTTTGT 360
QY      361 ATGACCAAGCTGGGTCCTGTATAGACACCTTCGACAACTGATGAGATTAAGTTT 420
DB      371 ATGACCAAGCTGGGTCCTGTATAGACACCTTCGACAACTGATGAGATTAAGTTT 420
QY      421 GACACCATATCTGAGAAAAACATCTGATGATCACTTTCTTTGCAAACTGAACTGC 480
DB      431 GACACCATATCTGAGAAAAACATCTGATGATCACTTTCTTTGCAAACTGAACTGC 480
QY      481 CGACTCTATCGAAAAACCAATCTCCAGATTAGTATCAGCAATGCGCTTTTGA 540
DB      491 CGACTCTATCGAAAAACCAATCTCCAGATTAGTATCAGCAATGCGCTTTTGA 540

```

```

DB      491 CGACTCTATCGAAAAACCAATCTCCAGATTAGTATCAGCAATGCGCTTTTGA 550
QY      541 GACAAATCCCTTACCTTCAATGAGACCTACAGAGACATGATGATGATATGAGACC 600
DB      551 GACAAATCCCTTACCTTCAATGAGACCTACAGAGACATGATGATGATATGAGACC 610
QY      601 AAGCTCAGCCCTTGAATTAAGGAAAAATGACAGCAATTCAGAGCCCATCAACAA 660
DB      611 AAGCTCAGCCCTTGAATTAAGGAAAAATGACAGCAATTCAGAGCCCATCAACAA 670
QY      661 TGGGTCGCAATTAAGCCAGGCGGATCCGATGATGATGATGATGATGATGATGAT 720
DB      671 TGGGTCGCAATTAAGCCAGGCGGATCCGATGATGATGATGATGATGATGATGAT 730
QY      721 GAGCTCACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 840
DB      731 GAGCTCACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 850
QY      781 TTGAGCCCTGAGAACCAAGAGAGAACTGTTCTACAAAGGCTGATGAGAGTCTTCA 840
DB      791 TTGAGCCCTGAGAACCAAGAGAGAACTGTTCTACAAAGGCTGATGAGAGTCTTCA 850
QY      841 GCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
DB      851 GCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 910
QY      901 GTGCTTGAAGTCCCTTCAAAAGTATGATGATGATGATGATGATGATGATGATGAT 960
DB      911 GTGCTTGAAGTCCCTTCAAAAGTATGATGATGATGATGATGATGATGATGATGAT 970
QY      961 GAGAGAGCTGAGGCAAGAGTGAAGAGAACTCAACCCCAAGGCTGTCAGAGTGTG 1020
DB      971 GAGAGAGCTGAGGCAAGAGTGAAGAGAACTCAACCCCAAGGCTGTCAGAGTGTG 1030
QY      1021 GATGAATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
DB      1031 GATGAATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1090
QY      1081 TTGAGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB      1091 TTGAGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1150
QY      1141 TCCAACTCCAGAGATTTGTCAGAGAGCCGAGATGATCTTATGTCAGATGATTC 1200
DB      1151 TCCAACTCCAGAGATTTGTCAGAGAGCCGAGATGATCTTATGTCAGATGATTC 1210
QY      1201 CATAGGCAATTTCTGAGGTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB      1211 CATAGGCAATTTCTGAGGTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1270
QY      1261 GTGATGTCGAGGCGTTCGCTTAAACCCCAAGGTCATTTCAAGGCAAGGCTTTTC 1320
DB      1271 GTGATGTCGAGGCGTTCGCTTAAACCCCAAGGTCATTTCAAGGCAAGGCTTTTC 1330
QY      1321 CTGCTTTTATAGAGAGATTCCTGTAACATTAATCTTCAAGGCAAGGCTTTTC 1380
DB      1331 CTGCTTTTATAGAGAGATTCCTGTAACATTAATCTTCAAGGCAAGGCTTTTC 1390
QY      1381 CTTGTGTAGTAA 1395
DB      1391 CTTGTGTAGTAA 1405

```

RESULT 3
ID AA229808 standard; DNA; 1525 BP.
AA229808;
AC AA229808;
XX 27-MAR-2000 (first entry)
DT 27-MAR-2000 (first entry)
XX Human antithrombin III DNA related sequence.
DB

OS	XX	Human antithrombin III; modified ATIII; elastase-resistant;
KW	XX	IGG activated neutrophil resistant; anti-thrombin activity; heparin;
KM	XX	anti-factor Xa activity; blood clotting disorder; sepsis; trauma; stroke;
KM	XX	thrombin activation-related pathological symptom; restenosis; thrombosis;
KM	XX	acute respiratory distress syndrome; thromboembolism; reocclusion; ds.
OS	XX	
OS	XX	Homo sapiens.
XX	XX	
XX	XX	W09958098-A2.
PN	XX	
PD	XX	18-NOV-1999.
XX	XX	
PF	XX	12-MAY-1999; 99WO-US10549.
XX	XX	
PR	XX	12-MAY-1999; 98US-0085197.
PR	XX	05-MAY-1999; 99US-0085197.
XX	XX	
PA	XX	(BOCK/) BOCK S C.
PA	XX	(PICA/) PICARD V.
XX	XX	(ZEND/) ZENDEHROUH P.
PI	XX	
XX	XX	Bock SC, Picard V, Zendeuhrouh P;
DR	XX	WPI: 2000-116274/10.
PT	XX	
PT	XX	New modified human antithrombin III compounds, used for treating e.g.
PT	XX	sepsis, trauma, acute respiratory distress syndrome, restenosis,
XX	XX	thrombosis, thromboembolism or stroke -
XX	XX	
PS	XX	Disclosure; Pages 71-72; 75pp; English.
XX	XX	
CC	XX	The present sequence is related to human antithrombin III (ATIII) DNA.
CC	XX	Modified ATIII have improved resistance to elastase and IGG-activated
CC	XX	neutrophils while retain anti-thrombin and anti-factor Xa activities.
CC	XX	They may be expressed as glycoforms with enhanced heparin affinity which
CC	XX	target the blood vessel wall more efficiently than ATIII with normal
CC	XX	heparin affinity. Modified ATIII can be used to treat thrombin
CC	XX	activation-related pathological symptoms due to sepsis, trauma, acute
CC	XX	respiratory distress syndrome, restenosis, thrombosis, thromboembolism
CC	XX	and stroke. It can also be used to reduce the risk of reocclusion
CC	XX	and restenosis in percutaneous transluminal coronary angioplasty,
CC	XX	thrombosis associated with surgery, ischaemia/reperfusion injury, and
CC	XX	coagulation abnormalities in cancer or surgical patients.
CC	XX	Note: There is no specific information about this sequence in the
CC	XX	specification.
XX	XX	
XX	XX	Sequence 1525 BP; 402 A; 385 C; 364 G; 374 T; 0 other;
XX	XX	
XX	XX	Query Match 99.8%; Score 1391.8; DB 21; Length 1525;
XX	XX	Best Local Similarity 99.9%; Pred No. 0;
XX	XX	Matches 1933; Conservative 0; Mismatches 2; Indels 0; Gaps 0
OY	1	ATGTAATTCATGATGATAGGAACCTGTAACTCTGGAAAAAGAGAGTTATCTTTGTTC 60
DB	47	ATGTAATTCATGATGATAGGAACCTGTAACTCTGGAAAAAGAGAGTTATCTTTGTTC 106
OY	61	TTGCGTCATATGGCTTCCTGGGACTCGCGGACCTGTACACGGAGCCCTGTGACATCTGC 120
DB	107	TTGCGTCATATGGCTTCCTGGGACTCGCGGACCTGTACACGGAGCCCTGTGACATCTGC 166
OY	121	ACAGCGAAGCCCGCGGACATTCCTCCATGATCCCATGTGTGATTCACGCTCCCGGAGAAG 180
DB	167	ACAGCGAAGCCCGCGGACATTCCTCCATGATCCCATGTGTGATTCACGCTCCCGGAGAAG 226
OY	181	AAGGCAACTGAGAGTGAAGGCTCAGAACAGAGATTCCTGGAGGCCACCAACCGGCTGTTC 240
DB	227	AAGGCAACTGAGAGTGAAGGCTCAGAACAGAGATTCCTGGAGGCCACCAACCGGCTGTTC 286
OY	241	TGGGAACGTGCCAAGGCCAATTCCTCGCTTGTACACATTTTATACAGACCTGGCAGAT 300
DB	287	TGGGAACGTGCCAAGGCCAATTCCTCGCTTGTACACATTTTATACAGACCTGGCAGAT 346
OY	301	TTCAAGAAATGACATGATTAACATTTTCTGTCTACCCCTGAGTATCTCCACGGCTTTTGTCT 360

Db	347	TCGAAGATGACAAATGATTAACATTTTCTGTGACCCCTGAGTATCTCCAGCGCTTTTGCT	406
Oy	361	ATGACCAAAGCTGGGTGGCTTGAATGACACCTCTCAGACACTGATGAGATTTAAATTT	420
Db	407	ATGACCAAAGCTGGGTGGCTTGAATGACACCTCTCAGACACTGATGAGATTTAAATTT	466
Oy	421	GACACCAATATCTGAGAAAACATCTGATCAGATCCACTTCTTTGGCAAACTGAATGTC	480
Db	467	GACACCAATATCTGAGAAAACATCTGATCAGATCCACTTCTTTGGCAAACTGAATGTC	526
Oy	481	CGACTCTATGAGAAAAGCCAAACAAATCTCTCCAGTTAGTATCAGCAATTCGCTTTTGG	540
Db	527	CGACTCTATGAGAAAAGCCAAACAAATCTCTCCAGTTAGTATCAGCAATTCGCTTTTGG	586
Oy	541	GACAAATCCCTTACCTTCAATGAGACCTCACAGGACATCAGTAGTGGTATATGAGACC	600
Db	587	GACAAATCCCTTACCTTCAATGAGACCTCACAGGACATCAGTAGTGGTATATGAGACC	646
Oy	601	AAGCTCAGCCCTCTGACCTTCAAGGAAAATGCAAGCAATTCAGAGCGGCCATCAACAA	660
Db	647	AAGCTCAGCCCTCTGACCTTCAAGGAAAATGCAAGCAATTCAGAGCGGCCATCAACAA	706
Oy	661	TGGGTGTTCCAATAGACCCGAAAGCCGAAATCACCATTCCTCTGGAAAGCATCAAT	720
Db	707	TGGGTGTTCCAATAGACCCGAAAGCCGAAATCACCATTCCTCTGGAAAGCATCAAT	766
Oy	721	GAGCTCACTGTCTTGGTGTGGTTAAACCACTTACTTCAAGGGCTGTGGAATGTCAA	780
Db	767	GAGCTCACTGTCTTGGTGTGGTTAAACCACTTACTTCAAGGGCTGTGGAATGTCAA	826
Oy	781	TTCAAGCCCTGAGAACACAAAGAGGAACTGTTCTACAAGGCTGATGGAAGTCTGTTC	840
Db	827	TTCAAGCCCTGAGAACACAAAGAGGAACTGTTCTACAAGGCTGATGGAAGTCTGTTC	886
Oy	841	GCATCTTATGATGTACAGAGAAAGGCAAGTCCGTTATCGGCGCGTGAAGGACCCAG	900
Db	887	GCATCTTATGATGTACAGAGAAAGGCAAGTCCGTTATCGGCGCGTGAAGGACCCAG	946
Oy	901	GTCCTGATGTTGCCCTTCAAGGATGATGACATCCATGTCCTCATCTTGGCCCAAGCT	960
Db	947	GTCCTGATGTTGCCCTTCAAGGATGATGACATCCATGTCCTCATCTTGGCCCAAGCT	1006
Oy	961	GAGAAAGACCTGGCGCAAGTGTGAGAAAGAACTCACCACAGAGTGTGCAAGAGATGGCTG	1020
Db	1007	GAGAAAGACCTGGCGCAAGTGTGAGAAAGAACTCACCACAGAGTGTGCAAGAGATGGCTG	1066
Oy	1021	GATAAATTTGAGAGATGATGCTGCTGTCACATGCCCCGCTTCCGCAATTGAGAGACGC	1080
Db	1067	GATAAATTTGAGAGATGATGCTGCTGTCACATGCCCCGCTTCCGCAATTGAGAGACGC	1126
Oy	1081	TTCACTTTGAAAGACGACGTGCAAGACATGAGGCTTGTGCATCTGTTCAAGCCCTGAAAG	1140
Db	1127	TTCACTTTGAAAGACGACGTGCAAGACATGAGGCTTGTGCATCTGTTCAAGCCCTGAAAG	1186
Oy	1141	TCCAAACTCCCAAGGATTTGTTGCGAAGAGCCGAGATGACCTTATGTCAGATGATTC	1200
Db	1187	TCCAAACTCCCAAGGATTTGTTGCGAAGAGCCGAGATGACCTTATGTCAGATGATTC	1246
Oy	1201	CATAAGGCATTTCTTGAAGTAAATGAAAGAGCGATGAGCAGCTGCAAGTACCGCTGTT	1260
Db	1247	CATAAGGCATTTCTTGAAGTAAATGAAAGAGCGATGAGCAGCTGCAAGTACCGCTGTT	1306
Oy	1261	GTCATTCCTGGCCGTTGCTGCTAAACCCCAACAGGGTGCTTTCAAGGGCAACAGGCTTTT	1320
Db	1307	GTCATTCCTGGCCGTTGCTGCTAAACCCCAACAGGGTGCTTTCAAGGGCAACAGGCTTTT	1366
Oy	1321	CTGGTTTTTAAAGAGAAGTTCCTCTGAACACTTATCTTTCATGAGGACAGATGACCAAC	1380
Db	1367	CTGGTTTTTAAAGAGAAGTTCCTCTGAACACTTATCTTTCATGAGGACAGATGACCAAC	1426
Oy	1381	CTTTGTGTAAAGTAA 1395	

DB 1427 CCTGTCTTAAGTAA 1441

RESULT 4
AAFS4046
ID AAF54046 standard; DNA; 1599 BP.
XX
XX AAF54046;
XX
XX 30-MAR-2001 (first entry)
XX
XX Human antithrombin III DNA, SEQ ID NO:43.
XX
XX Age-related gene regulation; gene expression; human protein C; hPC;
XX 5' UTR; 5' untranslated region; age-regulatable expression construct;
XX PEA-3 element; polyoma virus activator 3; antisense therapy;
XX gene therapy; thrombosis; cardiovascular disease; diabetes;
XX Alzheimer's disease; Parkinson's disease; cancer; osteoporosis;
XX osteoarthritis; dementia; ds.
XX
XX Homo sapiens.
XX
XX WO200075279-A2.
XX
XX 14-DEC-2000.
XX
XX 06-JUN-2000; 2000WO-US15728.
XX
XX 09-JUN-1999; 99US-0328925.
XX
XX (UNWI) UNIV MICHIGAN.
XX
XX Kurachi K, Kurachi S;
XX WPI; 2001-061708/07.
XX
XX New regulatory elements that control age-related gene expression,
XX useful in gene therapy and for reducing Factor IX expression -
XX
XX Disclosure: Fig 11; 225pp; English.
XX
XX The invention relates to nucleic acid sequences which regulate gene
XX expression in an age-related manner and/or in a liver-specific manner.
XX The invention identifies regions of the human factor IX (hFIX) gene, and
XX a region of the human protein C (hPC) gene, which are age-related
XX regulatory sequences. The hFIX age-related regulatory sequences are
XX designated AEs' (AAFS4016) and AEs' (AAFS4017) and are found in the 5'
XX UTR (at position 2164-2165 of AAF54018) and 3' UTR (at position
XX 34383-35655 of AAF54018) respectively. These elements act synergistically
XX to increase hFIX levels over the lifespan of an individual; however, they
XX can independently exert effects on hFIX mRNA in an age-related manner,
XX with AEs' acting to stabilize hFIX mRNA, and AEs' acting to increase hFIX
XX mRNA levels, over time. AEs' also directs liver-specific expression. The
XX hPC gene age-related regulatory sequence is found in the 5' UTR
XX (AAFS4081), and contains two PEA-3 (polyoma virus activator 3) elements
XX 5'-GAGGAA-3' and 5'-GAGGAA-3'. The age-related regulatory sequences of
XX the invention, along with their homologues, variants and fragments, may
XX be used in the construction of recombinant expression vectors for the
XX expression of a desired sequence in an age-related fashion in a host
XX cell. Preferred target genes for expression in such age-regulatable
XX expression vectors include those encoding proteins involved in blood
XX coagulation (e.g., the pro-coagulant factor IX and factor VIII, and the
XX anti-coagulant protein C and antithrombin III), human
XX alpha-1-antitrypsin, PEA-3 protein and reporter proteins such as
XX luciferase. Preferred promoters for use in such age-regulatable
XX expression vectors include the human factor IX promoter, the T7 promoter,
XX the T3 promoter and the SP6 promoter. The expression vectors of the
XX invention may be used in gene therapy to provide age-related and/or
XX liver-specific expression of target genes. Age-regulatable constructs may
XX be used in the treatment of such age-related conditions such as
XX thrombosis, cardiovascular disease, diabetes, Alzheimer's disease,
XX Parkinson's disease, cancer, osteoporosis, osteoarthritis and dementia.
XX Specifically, they may be used to express factor IX antisense mRNA in the

CC treatment of thrombotic conditions associated with the natural
CC age-related rise in factor IX expression. Transgenic cells or animals
CC that contain vectors of the invention are useful as models of these
CC diseases, in screening for potential therapeutic agents and for studying
CC normal processes such as aging and gene expression. Fragments and
CC homologues of age-related regulatory sequences, are useful as probes to
CC detect, isolate or identify other such sequences in samples. The present
CC sequence represents a nucleic acid sequence which may be incorporated
CC into a vector of the invention.
XX
XX Sequence 1599 BP; 421 A; 409 C; 375 G; 394 T; 0 other;
SQ
Query Match 99.8%; Score 1391.8; DB 22; Length 1599;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1393; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGTATTCCTCAATGTGATAGAACTGTAACTCTGGAAGAAAGAGTTATCTTTGTCC 60
DB 121 ATGTATTCCTCAATGTGATAGAACTGTAACTCTGGAAGAAAGAGTTATCTTTGTCC 180
QY 61 TTGCTGCTCAATGGCTCTGGGACTGGTGACTGTGACCTGTACCGGAGCCCTGTGAGATCTGC 120
DB 181 TTGCTGCTCAATGGCTCTGGGACTGGTGACTGTGACCTGTGAGAGCCCTGTGAGATCTGC 240
QY 121 ACAGCCAAAGCCGCGGACATTCCTCATGAATCCCATGTGATTCCTGCTCCCGAGAG 180
DB 241 ACAGCCAAAGCCGCGGACATTCCTCATGAATCCCATGTGATTCCTGCTCCCGAGAG 300
QY 181 AAGGCMACTGAGATGAGAGGCTCAGAACAGAAAGATCCCGAGGCGACCAACCGGCGTGC 240
DB 301 AAGGCMACTGAGATGAGAGGCTCAGAACAGAAAGATCCCGAGGCGACCAACCGGCGTGC 360
QY 241 TGGGAACTGTCAGAAAGGCAATTCCTGCTGCTCACTTCTATCAGACACTGCGCAT 420
DB 361 TGGGAACTGTCAGAAAGGCAATTCCTGCTGCTCACTTCTATCAGACACTGCGCAT 480
QY 301 TCCAGAAATGACATGATTAATCATTTCTGTCACCCCTGAGATCTCCAGCGCTTTGCT 360
DB 421 TCCAGAAATGACATGATTAATCATTTCTGTCACCCCTGAGATCTCCAGCGCTTTGCT 480
QY 361 ATGACCAAGCTGGGTGCTGTGATGACACCTTCAGCAACTGATGAGATTTAAGTTT 420
DB 481 ATGACCAAGCTGGGTGCTGTGATGACACCTTCAGCAACTGATGAGATTTAAGTTT 540
QY 421 GACACCATTTTGAGAAACATCTGATCATGATCCATCTTTTCCCAACTGAACTGC 480
DB 541 GACACCATTTTGAGAAACATCTGATCATGATCCATCTTTTCCCAACTGAACTGC 600
QY 481 CGACTCTATCGAAAGGCAACAAATCCCTCAAGTTAGATCAGCAATCGCTTTTGA 540
DB 601 CGACTCTATCGAAAGGCAACAAATCCCTCAAGTTAGATCAGCAATCGCTTTTGA 660
QY 541 GACAAATCCCTTACCTTCAATGAGACCTTACAGACATCAGTGAAGTTGATGAGACC 600
DB 661 GACAAATCCCTTACCTTCAATGAGACCTTACAGACATCAGTGAAGTTGATGAGACC 720
QY 601 AAGCTCCAGCCCTCGGACTTAAAGAAATATCAGAGCAATCAAGCGGCATCAACAA 660
DB 721 AAGCTCCAGCCCTCGGACTTAAAGAAATATCAGAGCAATCAAGCGGCATCAACAA 780
QY 661 TGGGTGTCATTAAGACCGAAGCGGATCAGCATGTCATTTCTCTCGAAGCCATCAAT 720
DB 781 TGGGTGTCATTAAGACCGAAGCGGATCAGCATGTCATTTCTCTCGAAGCCATCAAT 840
QY 721 GAGCTCACTGTTCTGCTGCTGTTAACAACATTTTCAAGGCGCTGTGGAAGTCAAG 780
DB 841 GAGCTCACTGTTCTGCTGCTGTTAACAACATTTTCAAGGCGCTGTGGAAGTCAAG 900
QY 781 TTGAGCCCTGGAACACAGAAAGAACTGTTTCAAGAGGCTGTGAGAGTCTGTGTTCA 840
DB 901 TTGAGCCCTGGAACACAGAAAGAACTGTTTCAAGAGGCTGTGAGAGTCTGTGTTCA 960
QY 841 GCATCTATGATGATCAGAAAGGCAAGTTCGTTATCGGCGCTGTGTAAGGACACCGAG 900

```

Db      961 GCATCTATATATACAGAGAGCAAGTCCGTTATCGGCGCTGAGAGCACCCAG 1020
Qy      901 GTGCTTGAATGAGCTTCAAGGTGATGACATCACCATGAGTCTCTATCTTGGCCAGCCT 960
Db      1021 GTGCTTGAATGAGCTTCAAGGTGATGACATCACCATGAGTCTCTATCTTGGCCAGCCT 1080
Qy      961 GAGAGAGCCTGCGCAAGGTGAGAGAGAACTCACCCCAAGGTGCTGAGAGTGGCTG 1020
Db      1081 GAGAGAGCCTGCGCAAGGTGAGAGAGAACTCACCCCAAGGTGCTGAGAGTGGCTG 1140
Qy      1021 GATGATTTGAGAGATGATGCTGAGTGTCCATATGCCCCGCTTCCGCAATTAGAGAGCGG 1080
Db      1141 GATGATTTGAGAGATGATGCTGAGTGTCCATATGCCCCGCTTCCGCAATTAGAGAGCGG 1200
Qy      1081 TTCAGTTTGAAGAGAGAGCTGCAAGACATGAGGCTTGTGATCTTTCAGCCCTGAAAAG 1140
Db      1201 TTCAGTTTGAAGAGAGAGCTGCAAGACATGAGGCTTGTGATCTTTCAGCCCTGAAAAG 1260
Qy      1141 TCCAAATCTCCAGGATTTGTTGCAAGAGCCGAGATGACCTCTATGCTTCAGATGCAATTC 1200
Db      1261 TCCAAATCTCCAGGATTTGTTGCAAGAGCCGAGATGACCTCTATGCTTCAGATGCAATTC 1320
Qy      1201 CATTAAGCAATTTCTTGAAGTAATGAAGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Db      1321 CATTAAGCAATTTCTTGAAGTAATGAAGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Qy      1261 GTGATTTGCTGCGCTTGCCTAAACCCCAACAGGAGTCTTTCAGAGCCCAACAGGCTTTC 1320
Db      1381 GTGATTTGCTGCGCTTGCCTAAACCCCAACAGGAGTCTTTCAGAGCCCAACAGGCTTTC 1440
Qy      1321 CTGGTTTATTAAGAAGAGTCTCTGACACATTAATCTTTCATGAGGAGAGAGAGAGAGAG 1380
Db      1441 CTGGTTTATTAAGAAGAGTCTCTGACACATTAATCTTTCATGAGGAGAGAGAGAGAGAG 1500
Qy      1381 CCTGTGTTAAGTAA 1395
Db      1501 CCTGTGTTAAGTAA 1515

```

RESULT 5

AAN30203

ID AAN30203 standard; cDNA; 1599 BP.

AC AAN30203;

DT 25-MAY-1992 (first entry)

DE Sequence encoding human antithrombin III (ATIII) from cDNA clones
DE pA62 and pA68.

KW Thrombosis; therapy; cardiovascular disorders; ss.

XX Homo sapiens.

FH Key location/Qualifiers

FT sig_peptide 121..216

FT mat_peptide 217..1515

FT polyA_signal 1599

FT /tag= c

FT /tag= a

FT /tag= b

FT /tag= c

FT /tag= c

FT /tag= c

FT /tag= c

FT /tag= c

```

XX      (GETH ) GENENTECH INC.
XX      Bock SC, Lawn RM;
XX      WPI: 1983-766797/38.
XX      P-PSDB; MAP30445.
XX      Human antithrombin III prod. by genetically modified cells - and
XX      corresp. cloning vehicles and expression vectors
XX      Disclosure; Fig 2; 14pp; English.
XX      pA62 and pA68 are overlapping cDNA clones for ATIII. pA62 is
XX      incomplete at the 3' end and pA68 comprises 400 bp from the 3'
XX      end, 84 bp 3' untranslated region and a poly(A) tail. Plasmids
XX      CC PAT111-87 and PAT111-J4 (claimed) contain a ligation product (pTA2)
XX      of pA62 and pA68 (see AAN30203).
XX      Sequence 1599 BP; 421 A; 408 C; 375 G; 395 T; 0 other;
SQ

```

Query Match

Best Local Similarity 99.7%; Score 1390.2; DB 4; Length 1599;

Matches 1392; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

Qy      1 ATGTAATTCATATGATATGAGAACTGTAACCTTGTGAAAAGAGAGCTTATCTTTTGTCC 60
Db      121 ATGTAATTCATATGATATGAGAACTGTAACCTTGTGAAAAGAGAGCTTATCTTTTGTCC 180
Qy      61 TTGCTGCTCATTTGGCTTCTGGGAGTGGGAGCTGACCTGTCAAGGAGGCTGTGAGACATCTGC 120
Db      181 TTGCTGCTCATTTGGCTTCTGGGAGTGGGAGCTGACCTGTCAAGGAGGCTGTGAGACATCTGC 240
Qy      121 ACAGCCAAAGCCGCGGAGACATTTCCATGAATCCATGTGCATTTTACCGCTCCCGAGAGAG 180
Db      241 ACAGCCAAAGCCGCGGAGACATTTCCATGAATCCATGTGCATTTTACCGCTCCCGAGAGAG 300
Qy      181 AAGGCACTGAGAGATGAGAGGCTCAGAAACAGAAAGATCCCGAGGCGCAACCAACCGGGGTGTC 240
Db      301 AAGGCACTGAGAGATGAGAGGCTCAGAAACAGAAAGATCCCGAGGCGCAACCAACCGGGGTGTC 360
Qy      241 TGGGAATGTCCTCAAGGCAATTCCTGCTTGTGCTACCATCTTGTATCAGACCTGCGACAT 300
Db      361 TGGGAATGTCCTCAAGGCAATTCCTGCTTGTGCTACCATCTTGTATCAGACCTGCGACAT 420
Qy      301 TCCAAAGATGACATGATTAACATTTTCTGTACACCCCTGAGTATCTCCACGCTTTTGTCT 360
Db      421 TCCAAAGATGACATGATTAACATTTTCTGTACACCCCTGAGTATCTCCACGCTTTTGTCT 480
Qy      361 ATGACCAAGTGGGAGCTGTGTAATGACACCTCCGACAGAACTGATGAGGATTTTAAGTTT 420
Db      481 ATGACCAAGTGGGAGCTGTGTAATGACACCTCCGACAGAACTGATGAGGATTTTAAGTTT 540
Qy      481 GACACCATATCTGAGAAAACATCTGATGAGATCTCACTTCTTTGCCAAATGTAAGTGC 480
Db      541 GACACCATATCTGAGAAAACATCTGATGAGATCTCACTTCTTTGCCAAATGTAAGTGC 600
Qy      481 CGACTCTATCGAAAAGCCAAATCTCTCAAGTTAGTATCAGCCAAATCGCCTTTTGTGA 540
Db      601 CGACTCTATCGAAAAGCCAAATCTCTCAAGTTAGTATCAGCCAAATCGCCTTTTGTGA 660
Qy      541 GACAAATCCCTTACCTTCAATGAGACTTACCAAGAGACATCAGAGTGTGATATGAGAGCC 600
Db      661 GACAAATCCCTTACCTTCAATGAGACTTACCAAGAGACATCAGAGTGTGATATGAGAGCC 720
Qy      601 AAGCTCCAGCCCTTGAAGCTTCAAGGAAAATGAGAGACATCCAGAGCGGCTATCAACAAA 660
Db      721 AAGCTCCAGCCCTTGAAGCTTCAAGGAAAATGAGAGACATCCAGAGCGGCTATCAACAAA 780
Qy      661 TGGGTGTCCAAATAGACCGAAGCGGAATCACCGATGTCATTTCCCTGGAGCCATCAAT 720
Db      781 TGGGTGTCCAAATAGACCGAAGCGGAATCACCGATGTCATTTCCCTGGAGCCATCAAT 840

```

Qy	721	GAGCTCACTGTTCCGTGTCCTGCTTAAACACACTTACTTCAAGGGCTGTGGAAGTCAAG	780
Db	841	GAGCTCACTGTTCCGTGTCCTGCTTAAACACACTTACTTCAAGGGCTGTGGAAGTCAAG	900
Qy	781	TTCAAGCCCTGAGAACACAAAGGAAGAACTGTCTTCAACAAGGCTGATGGAAGTCTGTTC	840
Db	901	TTCAAGCCCTGAGAACACAAAGGAAGAACTGTCTTCAACAAGGCTGATGGAAGTCTGTTC	960
Qy	841	GCATCTATGATGTACCCAGGAAGGCAAGTCCGTTATCGGCGCGTGTGGAAGGCAACCAG	900
Db	961	GCATCTATGATGTACCCAGGAAGGCAAGTCCGTTATCGGCGCGTGTGGAAGGCAACCAG	1020
Qy	901	GTGCTTGAGTGTCCCTTCMAAGGATGATACATCAACATGTCTCTATCTTCCCAAGCCT	960
Db	1021	GTGCTTGAGTGTCCCTTCMAAGGATGATACATCAACATGTCTCTATCTTCCCAAGCCT	1080
Qy	961	GAGAAAGAGCCTCGGCCAAGTGTGAGAAAGAACTCAACCCCAAGAGTGTCTGAGAGTGTG	1020
Db	1081	GAGAAAGAGCCTCGGCCAAGTGTGAGAAAGAACTCAACCCCAAGAGTGTCTGAGAGTGTG	1140
Qy	1021	GATGAAATTGAGAGAGATGATCTGTGTGTCCACATGCCCCCTTCCGATTTGAGAGACGC	1080
Db	1141	GATGAAATTGAGAGAGATGATCTGTGTGTCCACATGCCCCCTTCCGATTTGAGAGACGC	1200
Qy	1081	TTCAAGTTTGAAGAGACACGTGCAACATGTGGCCCTTGTGATCTGTTCAGCCCTGAAAG	1140
Db	1201	TTCAAGTTTGAAGAGACACGTGCAACATGTGGCCCTTGTGATCTGTTCAGCCCTGAAAG	1260
Qy	1141	TCCAAATCCCCAGGATATTGTTCAGAAAGCCGAGATGACCTCTAATGTCTCAGATGACATTC	1200
Db	1261	TCCAAATCCCCAGGATATTGTTCAGAAAGCCGAGATGACCTCTAATGTCTCAGATGACATTC	1320
Qy	1201	CATAAGGCATTCTTGAAGTAATGAGAAGGCACTGTAAGCAGCTGCAAGTACCGCTGTT	1260
Db	1321	CATAAGGCATTCTTGAAGTAATGAGAAGGCACTGTAAGCAGCTGCAAGTACCGCTGTT	1380
Qy	1261	GTGATGTGCGCCGTTGTGGCTAAACCCCAACAGGGGTGACTTTCAAGGGCCAAAGGCTTTC	1320
Db	1381	GTGATGTGCGCCGTTGTGGCTAAACCCCAACAGGGGTGACTTTCAAGGGCCAAAGGCTTTC	1440
Qy	1321	CTGGTTTTTATTAAGAGAATTCCTCTGGAACCTATTAATCTTCATGGGCAAGTACCAAC	1380
Db	1441	CTGGTTTTTATTAAGAGAATTCCTCTGGAACCTATTAATCTTCATGGGCAAGTACCAAC	1500
Qy	1381	CCTTGCTTAAAGTA 1395	
Db	1501	CTTGCTTAAAGTA 1515	

PN	EP568033-A.
XX	
PD	10-NOV-1993.
XX	
PF	08-APR-1993; 93EP-0105829.
XX	
PR	10-APR-1992; 92JP-0090488.
PR	22-FEB-1993; 93JP-0031855.
XX	
PA	(EISA) EISAI CO LTD.
XX	
PI	Kato H, Mizui Y, Nagaoka N, Seto T, Suzuki S;
PI	Yoshitake S;
XX	
DR	WPI; 1993-352985/45.
DR	P-PSDB; AAR42895.
XX	
PT	New human antithrombin III mutants of high antithrombin activity
PT	in absence of heparin - useful as anticoagulant for creating
PT	thrombotic disease
XX	
PS	Example 1; Page 35-37; 137pp; English.
XX	
CC	A commercially available human liver cDNA library was screened with
CC	a probe corresp. to amino acids 314-322 of AT III. Two positive
CC	overlapping clones were obtained. Their inserts were subcloned into
CC	pUC18 and the complete AT III coding sequence was obtained by
CC	ligating subcloned fragments. The full-length sequence provided the
CC	basis for a mutagenesis template; novel mutant AT III molecules were
CC	generated by site-directed techniques.
XX	
SO	Sequence 1395 BP; 368 A; 353 C; 343 G; 331 T; 0 other;

Accession	Gene	Protein	Location/Qualifiers
AA050415	AA050415 standard; CDNA; 1395 BP.		
XX			
AC	AA050415;		
XX			
DT	13-MAY-1994 (first entry)		
XX			
DE	Human antithrombin III cDNA.		
XX			
KW	AT III; mutant; serine protease inhibitor; serpin; thrombosis;		
KW	blood coagulation; anticonagulant; site-directed mutagenesis; ds.		
OS	Homo sapiens.		
XX			
FH	Key		
FT	CDS		1..1395
FT			/*tag= a
FT	sig_peptide		1..96
FT			/*tag= b
FT	mat_peptide		97..1392
FT			/*tag= C
FT			/note= "wild-type"

Query Match	Similarity	99.5%	Score	1388.6	DB	14	Length	1395
Best Local	Similarity	99.7%	Pred.	No.	0			
Matches	1391	Conservative	0	Mismatches	4	Indels	0	Gaps
QY	1	ATGATATTCGAAATGATAGAACTGTAACTCTGGAAAAAGAAAGTTATCTTTTGTC	60					
DB	1	ATGATATTCGAAATGATAGAACTGTAACTCTGGAAAAAGAAAGTTATCTTTTGTC	60					
QY	61	TTGCTGCTCATTTGGCTTTGGGACTGCGTGAACCTGTACCGGGAGCCCTGTGACATCTGC	120					
DB	61	TTGCTGCTCATTTGGCTTTGGGACTGCGTGAACCTGTACCGGGAGCCCTGTGACATCTGC	120					
QY	121	ACAGCGAAGCGCGGGGACATTCGCCATGAATCCCATGTGACATTTACCGCTCCCGGAGAG	180					
DB	121	ACAGCGAAGCGCGGGGACATTCGCCATGAATCCCATGTGACATTTACCGCTCCCGGAGAG	180					
QY	181	AAGGCAATGAGATGAGAGGCTCAGAAACGAAAGATCCCGAGGCCACCAACCGGCGTGC	240					
DB	181	AAGGCAATGAGATGAGAGGCTCAGAAACGAAAGATCCCGAGGCCACCAACCGGCGTGC	240					
QY	241	TGGGAATGTCGAAAGGCCAATTTCCGCTTTGCTACCACTTTCTATCAGACCTGGCAGAT	300					
DB	241	TGGGAATGTCGAAAGGCCAATTTCCGCTTTGCTACCACTTTCTATCAGACCTGGCAGAT	300					
QY	301	TCCAAAGATGACATGATATTAATTTTCTGTACACCCCGAGATATTCACAGGCTTTGCT	360					
DB	301	TCCAAAGATGACATGATATTAATTTTCTGTGTACACCCCGAGATATTCACAGGCTTTGCT	360					
QY	361	ATGACCAAGCTGGGTGCTCTGTATGACACCTTCGACGAACTGATGAGAGTATTTAAAGTTT	420					
DB	361	ATGACCAAGCTGGGTGCTCTGTATGACACCTTCGACGAACTGATGAGAGTATTTAAAGTTT	420					
QY	421	GACACCATATCTGAGAAAAACATCTGATCAGATCCACTTTCTTTTCCAAACTGACCTGC	480					
DB	421	GACACCATATCTGAGAAAAACATCTGATCAGATCCACTTTCTTTTCCAAACTGACCTGC	480					
QY	481	CGACTCTATGAAAAAGCCAAACAAATTCCTCCAGTTAGTATGAGCAATCGACTTTTGGG	540					
DB	481	CGACTCTATGAAAAAGCCAAACAAATTCCTCCAGTTAGTATGAGCAATCGACTTTTGGG	540					

QY 541 GACAAATCCCTTACCTTCAATGAGACCTTACAGACATCATGAGTTGGTATATGAGCC 600
 DB 541 GACAAATCCCTTACCTTCAATGAGACCTTACAGACATCATGAGTTGGTATATGAGCC 600
 QY 601 AAGCTCCAGCCCTTGAATCTTCAAGAGAAATGCAAGACATCCAGAGCCCATCAACAA 660
 DB 601 AAGCTCCAGCCCTTGAATCTTCAAGAGAAATGCAAGACATCCAGAGCCCATCAACAA 660
 QY 661 TGGGTGTCCATTAAGACCGAGGCCGATCACCGATGATTCCTCCGAGACCATCAAT 720
 DB 661 TGGGTGTCCATTAAGACCGAGGCCGATCACCGATGATTCCTCCGAGACCATCAAT 720
 QY 721 GAGCTCACTGTTCTGGTGTGTTAAACACATTTACTTCAAGGCCCTGGAAGTCAAG 780
 DB 721 GAGCTCACTGTTCTGGTGTGTTAAACACATTTACTTCAAGGCCCTGGAAGTCAAG 780
 QY 781 TTCAGCCCTGAGAACCAAGAGAAAGAACTGTTCTAAGGCTGATGAGAAAGTCGTCTCA 840
 DB 781 TTCAGCCCTGAGAACCAAGAGAAAGAACTGTTCTAAGGCTGATGAGAAAGTCGTCTCA 840
 QY 841 GCATCTATGATGATCAGAGAGCAAGTCCGTTATCGCGCGTGAAGCAACCCAG 900
 DB 841 GCATCTATGATGATCAGAGAGCAAGTCCGTTATCGCGCGTGAAGCAACCCAG 900
 QY 901 GTGCTTGAAGTTCCTTCAAGAGTGAATGACATCACATGAGTCTTCCCAAGCCT 960
 DB 901 GTGCTTGAAGTTCCTTCAAGAGTGAATGACATCACATGAGTCTTCCCAAGCCT 960
 QY 961 GAGAAAGCTTGGCCCAAGGTGAGAGAAAGAACTACCCCAAGGTGTGAGAGAGTGGC 1020
 DB 961 GAGAAAGCTTGGCCCAAGGTGAGAGAAAGAACTACCCCAAGGTGTGAGAGAGTGGC 1020
 QY 1021 GATGAATGAGAGAGATGATGCTGTGTCTCAATGCGCCGCTTCCGATTTGAGAGCGGC 1080
 DB 1021 GATGAATGAGAGAGATGATGCTGTGTCTCAATGCGCCGCTTCCGATTTGAGAGCGGC 1080
 QY 1081 TTCAGTTTGAAGAGAGAGCTGCAAGACATGAGCCTTGTGATGTTCAGCCCTGAAAG 1140
 DB 1081 TTCAGTTTGAAGAGAGAGCTGCAAGACATGAGCCTTGTGATGTTCAGCCCTGAAAG 1140
 QY 1141 TCCAACTCCCAAGGTGTTGTGAGAGAGCCGAGATGACCTCATGTCTCAGATGATTC 1200
 DB 1141 TCCAACTCCCAAGGTGTTGTGAGAGAGCCGAGATGACCTCATGTCTCAGATGATTC 1200
 QY 1201 CATTAAGCATTTCTTGAAGTAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
 DB 1201 CATTAAGCATTTCTTGAAGTAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
 QY 1261 GTGATTTGCTGGCCGTTGCTTAAACCCCAACAGGGTGAATTTCAAGGCCCAACAGGCC 1320
 DB 1261 GTGATTTGCTGGCCGTTGCTTAAACCCCAACAGGGTGAATTTCAAGGCCCAACAGGCC 1320
 QY 1321 CTGGTTTTTATAGAGAGAGTCTCTGAACATATTATCTTCAAGGCCAGAGTGAAGCAAG 1380
 DB 1321 CTGGTTTTTATAGAGAGAGTCTCTGAACATATTATCTTCAAGGCCAGAGTGAAGCAAG 1380
 QY 1381 CCTTGTGTTAAGTAA 1395
 DB 1381 CCTTGTGTTAAGTAA 1395

RESULT 7
 AAC83246
 ID AAC83246 standard; cDNA, 1444 BP.

AC AAC83246;
 XX
 DT 13-MAR-2001 (first entry)
 XX Human antithrombin III cDNA sequence.
 DE Human antithrombin III; antiangiogenic; angiogenesis inhibitor; anticoagulant;
 XX serine protease; human; ss.
 KW

XX Homo sapiens.
 OS
 XX
 PN W020069256-A1.
 XX
 FD 23-NOV-2000.
 XX
 PF 12-MAY-2000; 2000MO-US13052.
 XX
 PR 13-MAY-1999; 99US-0134174.
 XX
 PA (GENZ) GENZYME TRANSGENICS CORP.
 XX
 PI Meade H, Bourdon PR;
 XX
 DR WPI: 2001-024924/03.
 DR P-PSDB; AAB37963.
 PT Transgenically produced mutated human antithrombin III having
 PT antithrombin activity, useful for therapeutic and diagnostic
 PS applications -
 PS Disclosure; Fig 1; 45pp; English.

CC This invention relates to a transgenically produced mutated human
 CC antithrombin II polypeptide. Antithrombin III is a serine protease
 CC inhibitor which inhibits thrombin and activated forms of factors X, VII,
 CC IX, XI, and XII. Antithrombin III has antiangiogenic and anticoagulant
 CC activity. The mutated antithrombin III protein of the invention has uses
 CC in the therapy and diagnosis of disease and conditions involving
 CC angiogenesis. The invention includes methods for making a transgenic
 CC animal that produces the recombinant antithrombin III protein in its
 CC milk. The present sequence represents the human antithrombin III cDNA
 CC sequence.

SO Sequence 1444 BP; 377 A; 371 C; 353 G; 343 T; 0 other;

Query Match 99.4%; Score 1387; DB 22; Length 1444;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1390; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGTAATCCAAATGATGAGAACTGTAACCTGTGAAAGAGAGTTATCTTTGTC 60
 DB 16 ATCTATTCCAATGATGAGAACTGTAACCTGTGAAAGAGAGTTATCTTTGTC 75
 QY 61 TTGCTGCTCATTTGCTTGGAGCTGCTGACCTGTCAAGGAGCCCTGTGACATCTGC 120
 DB 76 TTGCTGCTCATTTGCTTGGAGCTGCTGACCTGTCAAGGAGCCCTGTGACATCTGC 135
 QY 121 ACAGCCAAAGCCGGGAGCAATTCCTCATGAAATCCCATGTGATTTACCGTCCCGAGAG 180
 DB 136 ACAGCCAAAGCCGGGAGCAATTCCTCATGAAATCCCATGTGATTTACCGTCCCGAGAG 195
 QY 181 AAGGCACTGAGAGATGAGAGGCTCAGAAACAGAGATCCCGAGAGCCCAACCCGGGTGC 240
 DB 196 AAGGCACTGAGAGATGAGAGGCTCAGAAACAGAGATCCCGAGAGCCCAACCCGGGTGC 255
 QY 241 TGGGAATCTCCAGGCCAATTCGCTTGTACCATCTTTATCAGACCTGCGAAT 300
 DB 256 TGGGAATCTCCAGGCCAATTCGCTTGTACCATCTTTATCAGACCTGCGAAT 315
 QY 301 TCCAAAGATGACAAATGATTAACATTTTCTGTGATACCCCTGAGATATCCAGGCTTTGCT 360
 DB 316 TCCAAAGATGACAAATGATTAACATTTTCTGTGATACCCCTGAGATATCCAGGCTTTGCT 375
 QY 361 ATGACCAAGCTGGGTGCTGTAATGACACCCCTCAGCAACTGATGAGGATTTAAGTT 420
 DB 376 ATGACCAAGCTGGGTGCTGTAATGACACCCCTCAGCAACTGATGAGGATTTAAGTT 435
 QY 421 GACACCATATCTGAGAAACATCTGATCAGATCACTTCTTTGCCAACTGAATCTGC 480
 DB 436 GACACCATATCTGAGAAACATCTGATCAGATCACTTCTTTGCCAACTGAATCTGC 495


```

QY 481 CGACTCTATCGAAGAAAGCAAAATCTCTCAAGTATGATACAGCAATCGCTTTTGA 540
DB 486 CGACTCTATCGAAGAAAGCAAAATCTCTCAAGTATGATACAGCAATCGCTTTTGA 555
QY 541 GACAAATCCCTTACCTTCAATGAGACTTACAGACATCATGAGTGTATATGAGGCC 600
DB 556 GACAAATCCCTTACCTTCAATGAGACTTACAGACATCATGAGTGTATATGAGGCC 615
QY 601 AAGCTCCAGCCCTTGAAGTCAAGAAATGAGAGCAATTCAGAGCCGCTTCAACAA 660
DB 616 AAGCTCCAGCCCTTGAAGTCAAGAAATGAGAGCAATTCAGAGCCGCTTCAACAA 675
QY 661 TGGGTGTCAATTAAGCCGAAAGCCGAATCAACGATGATCTTCCCTCGAAGCCATCAT 720
DB 676 TGGGTGTCAATTAAGCCGAAAGCCGAATCAACGATGATCTTCCCTCGAAGCCATCAT 735
QY 721 GAGCTCACTGTTCTGTGTGTGTTAACACCATTTTCAAGAGGCTGTGAAAGTCAAG 780
DB 736 GAGCTCACTGTTCTGTGTGTGTTAACACCATTTTCAAGAGGCTGTGAAAGTCAAG 795
QY 781 TTCAGCCCTGAGAAACAAAGAGAGCAATCTCTCAAGGCTGATGAGAGTCTGTCA 840
DB 796 TTCAGCCCTGAGAAACAAAGAGAGCAATCTCTCAAGGCTGATGAGAGTCTGTCA 855
QY 841 GCATCTATGATGTACAGAGAAAGCAAGTCCGTATTCGGCGGTGAGTGAAGCCAG 900
DB 856 GCATCTATGATGTACAGAGAAAGCAAGTCCGTATTCGGCGGTGAGTGAAGCCAG 915
QY 901 GTGCTTGAAGTCCCTTCAAGGTGATGATCATCATGATGCTTCAATCTTGGCCAAAGCT 960
DB 916 GTGCTTGAAGTCCCTTCAAGGTGATGATCATCATGATGCTTCAATCTTGGCCAAAGCT 975
QY 961 GAGAGAGCCCTGAGCAAGGTGAGAAAGCAATCCAGAGAGTCTGAGAGTGTGCTG 1020
DB 976 GAGAGAGCCCTGAGCAAGGTGAGAAAGCAATCCAGAGAGTCTGAGAGTGTGCTG 1035
QY 1021 GATGAATTTGAGAGATGATGCTGTGTGCTCATATGCCCCGCTTCCGATTTGAGAGCGC 1080
DB 1036 GATGAATTTGAGAGATGATGCTGTGTGCTCATATGCCCCGCTTCCGATTTGAGAGCGC 1095
QY 1081 TTCAGTTTGAAGAGCAAGCTGCAAGCATGAGGCTTGTGATCTGTTCAGCCCTGAAAG 1140
DB 1096 TTCAGTTTGAAGAGCAAGCTGCAAGCATGAGGCTTGTGATCTGTTCAGCCCTGAAAG 1155
QY 1141 TCCAAATCCCAAGATTTTTCAGAGAGCCGAGATGACCTTATGTCTCAAGATGATTC 1200
DB 1156 TCCAAATCCCAAGATTTTTCAGAGAGCCGAGATGACCTTATGTCTCAAGATGATTC 1215
QY 1201 CATAGGCAATTTCTTGAAGTAAATGAGAGGAGAGAGAGTCAAGTACCGCTGT 1260
DB 1216 CATAGGCAATTTCTTGAAGTAAATGAGAGGAGAGAGTCAAGTACCGCTGT 1275
QY 1261 GTGATTTGCTGGCGTTCGCTAAACCCCAACAGAGTGAATTTCAAGGCCCAACAGGCTTTC 1320
DB 1276 GTGATTTGCTGGCGTTCGCTAAACCCCAACAGAGTGAATTTCAAGGCCCAACAGGCTTTC 1335
QY 1321 CTGCTTTTATTAAGAGAGTCTCTGAAACATATTTATTTATGAGGAGAGTACCAAC 1380
DB 1336 CTGCTTTTATTAAGAGAGTCTCTGAAACATATTTATTTATGAGGAGAGTACCAAC 1395
QY 1381 CTTTGTGTTAAGTAA 1395
DB 1396 CTTTGTGTTAAGTAA 1410

```

RESULT 8
AAZ32160
ID AAZ32160 standard; cDNA; 1467 BP.

XX AAZ32160;
XX AC
XX 13-JAN-2000 (first entry)
XX

```

DE Human antithrombin III variant nucleotide sequence.
XX
XX Human; coding sequence polymorphism; vascular pathology gene;
XX polymorphic site; phenotype correlation; forensic; paternity testing;
XX medicine; genetic analysis; vascular disease; ds.
XX
XX Homo sapiens.
XX
XX MO9950454-A2.
XX
XX 07-OCT-1999.
XX
XX 26-MAR-1999; 99MO-US06473.
XX
XX 01-APR-1998; 98US-0054272.
XX
XX (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
XX Lander ES, Daley GQ, Cargill M, Ireland JS, Rozen SG;
XX
XX MPI; 1999-620066/53.
XX
XX P-PsDB; AAY49551.
XX
XX Determination of polymorphisms in genes, especially those identifying
XX predisposition to vascular disease -
XX
XX Claim 1; Fig 4; 134pp; English.
XX
XX AAZ32159 to AAZ32194 represent reference alleles for specifically
XX claimed nucleic acid sequences from the present invention which comprise
XX polymorphic sites as given in a table in the specification, selected
XX from 92 single nucleotide polymorphisms in which the nucleotide at the
XX polymorphic site is different from a nucleotide at the same site in a
XX reference allele. The nucleic acids, and primers and probes, are used to
XX identify polymorphisms, which may predispose an individual to disease,
XX especially a vascular disease. They can also be used in phenotype
XX correlations, forensic, paternity testing, medicine or genetic
XX analysis. AAY49550 to AAY49573 represent the proteins which correspond
XX to some of the reference alleles.
XX
XX Sequence 1467 BP; 361 A; 375 C; 364 G; 347 T; 0 other;
XX
XX Query Match 98.6%; Score 1375.6; DB 20; Length 1467;
XX Best Local Similarity 99.5%; Pred. No. 0;
XX Matches 1391; Conservative 0; Mismatches 4; Indels 3; Gaps 1;
XX
XX 1 ATGATTTCCAAATGTGATAGAACTGTAACTCTGGAAGAAAGAGATTATCTTTGTCC 60
XX 22 ATGATTTCCAAATGTGATAGAACTGTAACTCTGGAAGAAAGAGATTATCTTTGTCC 81
XX 61 TTGCTGCTCATTTGGCTTTTGGGACTGCTGACCTGTACCGGAGCCCTGTGACATCTGC 120
XX 82 TTGCTGCTCATTTGGCTTTTGGGACTGCTGACCTGTGACCGGAGCCCTGTGACATCTGC 141
XX 121 ACAGCCAGCCGCGGAGCAATTCCTCAATGATGATGATTTACCGCTCCCGGAGAG 180
XX 142 ACAGCCAGCCGCGGAGCAATTCCTCAATGATGATGATTTACCGCTCCCGGAGAG 201
XX 181 AAGGCACTGAGATGAGAGGCTCAGAACAGAAAGATCCCGAGGCGAC---CAACGCGGT 237
XX 202 AAGGCACTGAGATGAGAGGCTCAGAACAGAAAGATCCCGAGGCGACCAACAAACGCGGT 261
XX 238 GTCTGGGAACTGTCCAAAGCCCAATTCCTGCTTGTGTACCACTTTTATACAGACTGTGCA 297
XX 262 GTCTGGGAACTGTCCAAAGCCCAATTCCTGCTTGTGTACCACTTTTATACAGACTGTGCA 321
XX 298 GATTCCAGAAATGACAAATGATTAACATTTCTGTCACCCCTGAGATCTCCACGCTTTT 357
XX 322 GATTCCAGAAATGACAAATGATTAACATTTCTGTCACCCCTGAGATCTCTACGCTTTT 381
XX 358 GCTATGACCAAGCTGGGTGCTGTATGACACCTTCAGCAACTGTAGAGAGATTTAAG 417
XX 382 GCTATGACCAAGCTGGGTGCTGTATGACACCTTCAGCAACTGTAGAGAGATTTAAG 441

```

```

QY 418 TTGACACCATATCTGAGAAAACATCTGATCAGATCCACTCTTCTTGGCAAACTGAAAC 477
DB 442 TTGACACCATATCTGAGAAAACATCTGATCAGATCCACTCTTCTTGGCAAACTGAAAC 501
QY 478 TGGCGACTCTATGAGAAAAGCCAAATCTCTCAAGTATGATACACCAATGGCTTTT 537
DB 502 TGGCGACTCTATGAGAAAAGCCAAATCTCTCAAGTATGATACACCAATGGCTTTT 561
QY 538 GGAGACAAATCCCTTCAATGAGACCTACAGAGACATCAGTATGATATGGA 597
DB 562 GGAGACAAATCCCTTCAATGAGACCTACAGAGACATCAGTATGATATGGA 621
QY 598 GCCAAGCTCCAGCCCTGGAATTCAAGAAAATGAGAGAAATCCAGAGCCGATCAAC 657
DB 622 GCCAAGCTCCAGCCCTGGAATTCAAGAAAATGAGAGAAATCCAGAGCCGATCAAC 681
QY 658 AATGGGTGTCCATTAAGACCCGAAAGCCGAAATCAGGATGTCATTCCCTCGAAGCCATC 717
DB 682 AATGGGTGTCCATTAAGACCCGAAAGCCGAAATCAGGATGTCATTCCCTCGAAGCCATC 741
QY 718 AATGAGCTCACTGTGCTGGTGTGTTAACACCAATTAATCAAGGAGCCGTGGAAGTGA 777
DB 742 AATGAGCTCACTGTGCTGGTGTGTTAACACCAATTAATCAAGGAGCCGTGGAAGTGA 801
QY 778 AAGTTGAGCCCTGAGAACACAGAGAAAGAACTGTCTTACAGGCTGATGAGAGTGTGT 837
DB 802 AAGTTGAGCCCTGAGAACACAGAGAAAGAACTGTCTTACAGGCTGATGAGAGTGTGT 861
QY 838 TCAGCATCTATATGATGATACAGAGAAAGCAAGTTCCTGTTATGCGGCTGCTGAAGCAC 897
DB 862 TCAGCATCTATATGATGATACAGAGAAAGCAAGTTCCTGTTATGCGGCTGCTGAAGCAC 921
QY 898 CAGGTCTGTGAGTGTGCTTCAAGGTGATGATCATACCATGCTCTCATCTTGCCCAAG 957
DB 922 CAGGTCTGTGAGTGTGCTTCAAGGTGATGATCATACCATGCTCTCATCTTGCCCAAG 981
QY 958 CCTGAGAGAGCCTGAGCAAGGTGAGAGAAAGAACTGACCCAGAGGTGCTGAGAGTGTG 1017
DB 982 CCTGAGAGAGCCTGAGCAAGGTGAGAGAAAGAACTGACCCAGAGGTGCTGAGAGTGTG 1041
QY 1018 CTGAGATGATGAGAGAGATGATGCTGTGTGCTCAATGCGGCTTCGCAATGAGAGAC 1077
DB 1042 CTGAGATGATGAGAGAGATGATGCTGTGTGCTCAATGCGGCTTCGCAATGAGAGAC 1101
QY 1078 GGGTTCAGTTTGAAGAGACAGTGAAGACATGGGCTTGTGATCTGTTCAAGCCCTGAA 1137
DB 1102 GGGTTCAGTTTGAAGAGACAGTGAAGACATGGGCTTGTGATCTGTTCAAGCCCTGAA 1161
QY 1138 AAGTCCAAACTCCAGGATATGTTGCAAGAGCCGAGATGACCTTATGTTCTCAGATGCA 1197
DB 1162 AAGTCCAAACTCCAGGATATGTTGCAAGAGCCGAGATGACCTTATGTTCTCAGATGCA 1221
QY 1198 TTCCATTAAGGCAATTTCTGAGTAATGTAAGAGAGCAAGCACTGCAATGACCGT 1257
DB 1222 TTCCATTAAGGCAATTTCTGAGTAATGTAAGAGAGCAAGCACTGCAATGACCGT 1281
QY 1258 GTTGTATGCTGGCCGTTGCTGAACCCCAAGAGGTGACTTCAAGGCCCAAGAGCCCT 1317
DB 1282 GTTGTATGCTGGCCGTTGCTGAACCCCAAGAGGTGACTTCAAGGCCCAAGAGCCCT 1341
QY 1318 TTCTGTTTATTAAGAGAGAGTTCTCTGAAACATATTAATCTTCAATGGGAGAGTATGCC 1377
DB 1342 TTCTGTTTATTAAGAGAGAGTTCTCTGAAACATATTAATCTTCAATGGGAGAGTATGCC 1401
QY 1378 AACCTGTGTAGTAA 1395
DB 1402 AACCTGTGTAGTAA 1419

```

RESULT 9
AAV41731
ID AAV41731 standard; DNA; 1382 BP.

```

XX AC AAV41731;
XX DT 20-NOV-1998 (first entry)
XX DE Codon-optimised Ramy3D signal fused to DNA encoding mature AtIII.
XX KW Protein expression; monocotyledon plant cell;
XX KW glycosylated alpha 1-antitrypsin; AAT; glycosylated antithrombin III;
XX KW AtIII; human serum albumin; HSA; subtilisin BPN'; treatment; emphysema;
XX KW antithrombotic; blood replacement; ss.
XX OS Synthetic.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT misc_feature 1..75
FT /tag= a
FT /note= "codon-optimised Ramy3D signal sequence"
FT /tag= b
FT /note= "encodes mature AtIII"
XX PN MO9836085-A1.
XX PD 20-AUG-1998.
XX PF 13-FEB-1998; 98WO-US03068.
XX PR 13-FEB-1997; 97US-0038170.
XX PR 13-FEB-1997; 97US-0037991.
XX PR 13-FEB-1997; 97US-0038168.
XX PR 13-FEB-1997; 97US-0038169.
XX PA (PHYT-) APPLIED PHYTOLOGICS INC.
XX PI Rodriguez RL, Sutcliffe TD;
XX DR WPI; 1998-467179/40.
XX PT Expressing mature, glycosylated proteins in monocotyledonous plant
XX PT cells - from chimeric gene including signal peptide sequence,
XX PT specifically therapeutic agents and industrial enzymes
XX PS Disclosure; Page 34 iv; 53pp; English.
XX XX
XX CC The present sequence encodes a fusion protein of codon-optimised Ramy3D
XX CC signal sequence/mature antithrombin III (AtIII). The protein is used
XX CC to exemplify the invention. The specification describes a method for
XX CC producing mature heterologous protein in monocotyledonous plant cells.
XX CC The method comprises transforming the cells with a chimeric gene
XX CC comprising a monocotyledon transcription regulator, inducible either
XX CC during seed maturation or by adding/removing a small molecule, DNA
XX CC encoding the heterologous protein, and DNA encoding a signal peptide,
XX CC with the signal peptide causing secretion of the protein from the cell.
XX CC Proteins expressed in this manner include mature glycosylated alpha
XX CC 1-antitrypsin (AAT) with a glycosylation pattern that significantly
XX CC increases its serum half-life, mature glycosylated antithrombin III
XX CC (AtIII), mature human serum albumin (HSA) having the native folding
XX CC pattern as shown by bilirubin-binding characteristics, or mature active
XX CC subtilisin BPN'. These proteins are useful therapeutically (e.g. AAT for
XX CC treating emphysema, AtIII as antithrombotic and HSA as blood replacement)
XX CC or as industrial enzymes (BPN' is used in detergents).
XX SQ Sequence 1382 BP; 358 A; 370 C; 339 G; 315 T; 0 other;

Query Match 92.9%; Score 1295.6; DB 19; Length 1382;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```

QY 138 CATTCATGATGATCCCATGATGATTTACCGCTCCCGGAGAGAAAGGCAACTGAGATGA 197
 DB 117 CATTCATGATGATCCCATGATGATTTACCGCTCCCGGAGAGAAAGGCAACTGAGATGA 176
 QY 198 GGGCTCAGAAAGAAAGATCCCGAGGCGCAACACCGGCGTGTCTGGAACTGTCCAAAGC 257
 DB 177 GGGCTCAGAAAGAAAGATCCCGAGGCGCAACACCGGCGTGTCTGGAACTGTCCAAAGC 236
 QY 258 CAATTCGCGCTTGTCTACCACTTTCTATCAGCACTGAGAGATTCAGAAATGACATGA 317
 DB 237 CAATTCGCGCTTGTCTACCACTTTCTATCAGCACTGAGAGATTCAGAAATGACATGA 296
 QY 318 TAACATTTCTGTCAACCCCTGAGTATCTCAACGCGCTTTGCTATGACCAAGCTGGGTC 377
 DB 297 TAACATTTCTGTCAACCCCTGAGTATCTCAACGCGCTTTGCTATGACCAAGCTGGGTC 356
 QY 378 CTGTATGACACCCCTCAGCAACCTGATGAGGATATTAAATTGAACCATATCTGAGA 437
 DB 357 CTGTATGACACCCCTCAGCAACCTGATGAGGATATTAAATTGAACCATATCTGAGA 416
 QY 438 AACATCTGATGATTCATCTTCTTTGCGCAACTGAACCTGCGCACTTATCGAAAGC 497
 DB 417 AACATCTGATGATTCATCTTCTTTGCGCAACTGAACCTGCGCACTTATCGAAAGC 476
 QY 498 CAACAATCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 557
 DB 477 CAACAATCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 536
 QY 558 CAATGAGACCTACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 617
 DB 537 CAATGAGACCTACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 596
 QY 618 CTTCAAGGAAATGACAGACATCCAGAGCGGCATCAACAATGAGTGTCCAAATGAC 677
 DB 597 CTTCAAGGAAATGACAGACATCCAGAGCGGCATCAACAATGAGTGTCCAAATGAC 656
 QY 678 CGAAGCGCAATCACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 737
 DB 657 CGAAGCGCAATCACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 716
 QY 738 GCTGTAAACACATTTACTTCAAGGCGCTGTGAAAGTCAAGTCCAGCTGAGAACAC 797
 DB 717 GCTGTAAACACATTTACTTCAAGGCGCTGTGAAAGTCAAGTCCAGCTGAGAACAC 776
 QY 798 AAGGAAGAACTGTTCTTCAAGGCTGATGAGAGTGTGTTCAGCATCTATGATGATCA 857
 DB 777 AAGGAAGAACTGTTCTTCAAGGCTGATGAGAGTGTGTTCAGCATCTATGATGATCA 836
 QY 858 GGAAGGCAAGTCCGTTATCGGCGGCTGTAAGGCAACCGGCTTGAAGTCCCTT 917
 DB 837 GGAAGGCAAGTCCGTTATCGGCGGCTGTAAGGCAACCGGCTTGAAGTCCCTT 896
 QY 918 CAAAGGTATGACATCACCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 977
 DB 897 CAAAGGTATGACATCACCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 956
 QY 978 GGTGGAAGAACTACACCCAGAGGTGCTGAGAGATGCTGAGATGATGAGAGAT 1037
 DB 957 GGTGGAAGAACTACACCCAGAGGTGCTGAGAGATGCTGAGATGATGAGAGAT 1016
 QY 1038 GATGCGTGTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1097
 DB 1017 GATGCGTGTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1076
 QY 1098 GGTGCAAGCATGAGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1157
 DB 1077 GGTGCAAGCATGAGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1136
 QY 1158 TGTGCAAGGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1217
 DB 1137 TGTGCAAGGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1196

QY 1218 GGTAAATGAAAGGAGATGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 1277
 DB 1197 GGTAAATGAAAGGAGATGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 1256
 QY 1278 GGTAAATGAAAGGAGATGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 1337
 DB 1257 GGTAAATGAAAGGAGATGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 1316
 QY 1338 AGTTCCTGAAACCTATATCTTCAATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 1395
 DB 1317 AGTTCCTGAAACCTATATCTTCAATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 1374
 RESULT 10
 ID AAV41727 standard; DNA; 1299 BP.
 AC AAV41727;
 XX
 DT 20-NOV-1998 (first entry)
 XX
 DE DNA encoding the mature protein of antithrombin III (AIII).
 XX
 KW Protein expression; monocytledon plant cell;
 KW glycosylated alpha 1-antitrypsin; AAT; glycosylated antithrombin III;
 KW AIII; human serum albumin; HSA; subtilisin BPN'; treatment; emphysema;
 KW antithrombotic; blood replacement; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9836085-A1.
 XX
 PD 20-AUG-1998.
 XX
 PF 13-FEB-1998; 98WO-US03068.
 XX
 PR 13-FEB-1997; 97US-0038170.
 XX
 PR 13-FEB-1997; 97US-0037991.
 PR 13-FEB-1997; 97US-0038168.
 PR 13-FEB-1997; 97US-0038169.
 XX
 PA (PHYT-) APPLIED PHYTOLOGICS INC.
 XX
 PI Rodriguez RL, Sutcliffe TD;
 XX
 DR WPI; 1998-467179/40.
 XX
 PT P-PSDB; AAM59840.
 XX
 PS
 XX
 PT Expressing mature, glycosylated proteins in monocytledonous plant
 PT cells - from chimeric gene including signal peptide sequence,
 PT specifically therapeutic agents and industrial enzymes
 XX
 PS Disclosure; Page 31; 53pp; English.
 XX
 CC The present sequence encodes the mature protein of antithrombin III
 CC (AIII). The protein is used to exemplify the invention. The
 CC specification describes a method for producing mature heterologous
 CC protein in monocytledonous plant cells. The method comprises
 CC transforming the cells with a chimeric gene comprising a monocytledon
 CC transcription regulator, inducible either during seed maturation or by
 CC adding/removing a small molecule, DNA encoding the heterologous protein,
 CC and DNA encoding a signal peptide, with the signal peptide causing
 CC secretion of the protein from the cell. Proteins expressed in this
 CC manner include mature glycosylated alpha 1-antitrypsin (AAT) with a
 CC glycosylation pattern that significantly increases its serum half-life,
 CC mature glycosylated antithrombin III (AIII), mature human serum albumin
 CC (HSA) having the native folding pattern as shown by bilirubin-binding
 CC characteristics, or mature active subtilisin BPN'. These proteins are
 CC useful therapeutically (e.g. AAT for treating emphysema, AIII as
 CC antithrombotic and HSA as blood replacement) or as industrial enzymes
 CC (BPN' is used in detergents).
 XX
 SQ Sequence 1299 BP; 346 A; 334 C; 321 G; 298 T; 0 other;

Query Match 92.8%; Score 1294.2; DB 19; Length 1299;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1296; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 97 CAGGGAGCCCTGTGACATCTGCACAGCCAGCCGGGAGCATCCCATGATCCCATG 156
   |||
Db 1 CAGGAGAGCCCTGTGACATCTGCACAGCCAGCCGGGAGCATCCCATGATCCCATG 60
   |||

QY 157 TGCATTACCGCTCCCGGAGAGAGAGCACTGAGATGAGGGCTCAGAAACAGAAATC 216
   |||
Db 61 TGCATTACCGCTCCCGGAGAGAGAGCACTGAGATGAGGGCTCAGAAACAGAAATC 120
   |||

QY 217 CCGGAGGCCACCAACCGGCTGTCTGTGGAATCTGCCAGGCCAATTCGCTTGTCTACC 276
   |||
Db 121 CCGGAGGCCACCAACCGGCTGTCTGTGGAATCTGCCAGGCCAATTCGCTTGTCTACC 180
   |||

QY 277 ACTTTCATGAGACCGGAGATTCACAAATGATGACAAATGATTAACATTTTCTGTCAACC 336
   |||
Db 181 ACTTTCATGAGACCGGAGATTCACAAATGATGACAAATGATTAACATTTTCTGTCAACC 240
   |||

QY 337 CTGAGTATCTCAACGGCTTTTGTATGACCAAGCTGGGTGCTGTATGACACCTTCAG 396
   |||
Db 241 CTGAGTATCTCAACGGCTTTTGTATGACCAAGCTGGGTGCTGTATGACACCTTCAG 300
   |||

QY 397 CAATGATGAGGATATTAAAGTTTGACACCAATCTGAGAAAACATCTGATCATTCAC 456
   |||
Db 301 CAATGATGAGGATATTAAAGTTTGACACCAATCTGAGAAAACATCTGATCATTCAC 360
   |||

QY 457 TTCTCTTTGCGCAACCTGAACTGCGGCTCTATGAAAACCAAACTCTTCATGATTA 516
   |||
Db 361 TTCTCTTTGCGCAACCTGAACTGCGGCTCTATGAAAACCAAACTCTTCATGATTA 420
   |||

QY 517 GTATCAGCCCAATGCGCTTTTGTGAGACAAATCCCTTAACCTTCAATGAGACTTCA 576
   |||
Db 421 GTATCAGCCCAATGCGCTTTTGTGAGACAAATCCCTTAACCTTCAATGAGACTTCA 480
   |||

QY 577 ATGATGAGATTTGTATATGAGCCCAAGCTCCAGCCCTGAGACTTCAAGAAAATGCA 636
   |||
Db 481 ATGATGAGATTTGTATATGAGCCCAAGCTCCAGCCCTGAGACTTCAAGAAAATGCA 540
   |||

QY 637 CAATCAGAGCGGCTCAACCAATGAGGTGTCCAAATAAGCGGAGGCGCAATCACCGAT 696
   |||
Db 541 CAATCAGAGCGGCTCAACCAATGAGGTGTCCAAATAAGCGGAGGCGCAATCACCGAT 600
   |||

QY 697 GTCAATTCCTCGGAGGCAATGAGCTCACTGTCTGTGTGTTTAAACACATTTAC 756
   |||
Db 601 GTCAATTCCTCGGAGGCAATGAGCTCACTGTCTGTGTGTTTAAACACATTTAC 660
   |||

QY 757 TTCAAGGCGCTGTGGAAGTCAAAAGTTTCAAGCTGAGAAACAAGAAAGAACTGTCTAC 816
   |||
Db 661 TTCAAGGCGCTGTGGAAGTCAAAAGTTTCAAGCTGAGAAACAAGAAAGAACTGTCTAC 720
   |||

QY 817 AAGGCTGATGAGAGTGTGTGTCAGCATCTATGATGACAGAGAGGCAAGTCCGTTAT 876
   |||
Db 721 AAGGCTGATGAGAGTGTGTGTCAGCATCTATGATGACAGAGAGGCAAGTCCGTTAT 780
   |||

QY 877 CGGCGGCTGTGAGAGGCAACAGGTGCTTGAAGTTCCTTCAAGGTGATGACATCAC 936
   |||
Db 781 CGGCGGCTGTGAGAGGCAACAGGTGCTTGAAGTTCCTTCAAGGTGATGACATCAC 840
   |||

QY 937 ATGCTCTCATCTTTGCCCAAGCTGAGAAAGAGCTGCGCAAGGTGAGAAAGAACTCAC 996
   |||
Db 841 ATGCTCTCATCTTTGCCCAAGCTGAGAAAGAGCTGCGCAAGGTGAGAAAGAACTCAC 900
   |||

QY 997 CAGAGGTGTGAGAGAGTGTGTGATGATGAGAGAGATGATGCTGTGTCTCCATG 1056
   |||
Db 901 CAGAGGTGTGAGAGAGTGTGTGATGATGAGAGAGATGATGCTGTGTCTCCATG 960
   |||

QY 1057 CCGCGCTTCCGATTTGAGAGCGGCTTCAAGTTTGAAGAGAGAGCTGCAAGCATGGGCTT 1116
   |||
Db 961 CCGCGCTTCCGATTTGAGAGCGGCTTCAAGTTTGAAGAGAGAGCTGCAAGCATGGGCTT 1020
   |||

```

```

QY 1117 GTGATCTGTAGGCTTGAAAAGTCAAACTCCAGGATATTTGTGAGAAAGCCGAGAT 1176
   |||
Db 1021 GTGATCTGTAGGCTTGAAAAGTCAAACTCCAGGATATTTGTGAGAAAGCCGAGAT 1080
   |||

QY 1177 GACCTCTATGTCTCAGATGATTCATTAAGGATTTCTTGAAGTAAAGAAAGGACAGT 1236
   |||
Db 1081 GACCTCTATGTCTCAGATGATTCATTAAGGATTTCTTGAAGTAAAGAAAGGACAGT 1140
   |||

QY 1237 GAAGAGCTGCAAGTACCGCTGTGTGATGTGCTGCGCTTAACCCCAACAGGGTG 1296
   |||
Db 1141 GAAGAGCTGCAAGTACCGCTGTGTGATGTGCTGCGCTTAACCCCAACAGGGTG 1200
   |||

QY 1297 ACTTTCAGGCGCAAGGCTTCTGTGTTTAAAGAGAGTCTCTGAACCTAAT 1356
   |||
Db 1201 ACTTTCAGGCGCAAGGCTTCTGTGTTTAAAGAGAGTCTCTGAACCTAAT 1260
   |||

QY 1357 ATCTTCATGGGAGAGTACCAACCTTGTGTTAAGTAA 1395
   |||
Db 1261 ATCTTCATGGGAGAGTACCAACCTTGTGTTAAGTAA 1299
   |||

RESULT 11
AAS83474
ID AAS83474 standard; cDNA; 2123 BP.
XX
AC AAS83474;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #19278.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW Food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN MO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI, 2001-639362/73.
DR P-PDB; ABG19287.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 19278; 103bp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations

```


PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 1; SEQ ID No 19279; 103bp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. A564197-A594564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 1724 BP; 430 A; 439 C; 418 G; 437 T; 0 other;
 Query Match 36.7%; Score 512.4; DB 23; Length 1724;
 Best Local Similarity 79.9%; Pred. No. 9.2e-136;
 Matches 1195; Conservative 0; Mismatches 196; Indels 105; Gaps 46;

DB 675 TTAGTATAGCCCAATGCGCTTTTGGAGACAATTCCTTACTTCATGAGACCTACGAG 734
 QY 573 GGACATCAGTAGGTTGTATATGAGCCAGCTCCAGCCC--TGACCTTCAAGAAA 629
 DB 735 GGACATCAGTAGGTTGTATATGAGCCAGCTCCAGCCCCTGGACCTTCAAGGAAA 794
 QY 630 TGCAGAGCAAT--CCAGAGCGGCATCAACAAT--GGGTGTCATATAGACCGAAGGCG 686
 DB 795 TGCAGAGCAATTCAGAGCGGCATCAACAATGGGGTGTCCATAAAGCCGAAGGCG 854
 QY 687 AA-TCAACGATGTCATTCCTCGAAG--CCATCAATAGCTCAGCTTCTG--GT 737
 DB 855 AATTACCGAGTGTATTTTCCCTGGGAAGGCATTCATATAGCTCATGTTCTGGGTG 914
 QY 738 GCTGTATACACCATTTACTTCAAGG--CCTGTGAAGTCAAGTTCAAGCCTTGA 793
 DB 915 TGGGTTTAAACCATTTACTTCAAGGGCGCTGTGGAAGGTCAAAAGTTCAAGCCTGAGA 974
 QY 794 ACACAGAGAGAAATGTTTCTACAAAGCTGAT--GGAGAGTGTG--TTGAGATCTATG-- 849
 DB 975 ACACAGAGAGAAATGTTTCTACAAAGCTGATGGAGAGTGTGTTTACAGATCTATGGA 1034
 QY 850 -ATGACAGAGAGAGCAAGTCCGTTAT--CGAGCGTGTGTAAGGACCCAGGTGCT 906
 DB 1035 TGTTACAGAGAGAGCAAGTCCGTTATTCGGGCGCGGTGCTTGAAGGACCCGGTCTT 1094
 QY 907 --GAGTTCCTTCAAAAGGTGATGATCA--CCATGCTCTCATCTTCCCAAGCC-- 959
 DB 1095 GATGTTTCCCTTCAAAAGGTGATGATCA--CCATGCTCTCATCTTCCCAAGCCCTG 1154
 QY 960 TGAGAGAGAGCTGGGCAAGGTGAGAA--GSAATCAACCCAGAGGTGCT--GCAGAGGTG 1016
 DB 1155 AAGAGAGAGCTGGGCAAGGTGAGAAAGGAACTACCCAGAGGTGCTTCCAAAGAGT 1214
 QY 1017 GCTGG--ATGAATTGAGAGAGATGATGCTGTGATGCTCAGATGCCCCGCTT--CCGAT 1070
 DB 1215 GCTGGAGATGATTTTGAAGAGAGATGATGCTGTGATGCTCAGATGCCCCGCTTCCGAT 1274
 QY 1071 TGAGAGAGAGCTTCAAGTTGAAGAGAGAGCTGCAAGACATGGGCTTGTGATCTTCAAG 1130
 DB 1275 TGAGAGAGAGCTTCAAGTTGAAGAGAGAGAGCTGCAAGACATGGGAGCTTGTGATG 1334
 QY 1131 CCTGAAAGTCC-----AAATCCAGAGATTTGTTGAGAAAG--CG 1172
 DB 1335 TCTGTGAGGCCCCCTGAAAGATCCCAAACTCCCCGGATTTGTGCAAGAGCCCG 1394
 QY 1173 AGATGACCTCTATGCTCAGATG--ATTCTAAGAGCATTT--CTTGAAGTAATGA 1226
 DB 1395 AGATGACCTCTATGCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1454
 QY 1227 AGAGAGAGAGAGAGAG--CTGCAAGTA--CCGTGTGTGATG--CTGAGCCGTTG--CTAA 1282
 DB 1455 AGAGAGAGAGAGAGAGAGCTGCAAGATCCCGTGTGTGATGCTGAGCCGTTCCCTAA 1514
 QY 1283 ACCCAAGAGAGAGAG--TTTCAAGGCAAGAGGCTTCTCT--GGTTTTTAAAGAGAG 1339
 DB 1515 ACCCAAGAGAGAGAGAGCTTCAAGGCAAGAGGCTTCTCTCTGAGTTTAAAGAGAG 1574
 QY 1340 TTCTCTC-TGAACATATTATCTTCAATGGGAGAG--TAGCCAACTTGTGTTAA 1391
 DB 1575 TTCTCTTGAACATATTATCTTCAATGGGAGAGAGTTAGCAACCTTGTGTTAA 1630

RESULT 13
 ABA63081
 ID ABA63081 standard; DNA; 578 BP.
 XX
 AC ABA63081;
 XX
 DT 01-FEB-2002 (first entry)
 XX
 DE Human foetal liver single exon nucleic acid probe #11386.
 XX

Db 241 GTGCTGCAAGAGGCTGATGATGAGAGATATCTGTGTGTCACATGCCCCC 300
 QY 1063 TTCCGATTGAGAGCGCTTCAGTTGAGAGACAGCTGCAAGACATGGGCTTGCGAT 1122
 Db 301 TTCCGATTGAGAGCGCTTCAGTTGAGAGACAGCTGCAAGACATGGGCTTGCGAT 360
 QY 1123 CTGTTGAGCCCTGAAAAAGTCCAAACTCCAGGTAT 1157
 Db 361 CTGTTGAGCCCTGAAAAAGTCCAAACTCCAGGTAT 395
 RESULT 15
 AAS83471
 ID AAS83471 standard; cDNA; 433 BP.
 XX
 AC AAS83471;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #19275.
 XX
 KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR P-PSDB; ABG19284.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 1; SEQ ID No 19275; 103bp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 433 BP; 113 A; 111 C; 108 G; 100 T; 1 other;

Query Match 24.4%; Score 340; DB 23; Length 433;
 Best Local Similarity 100.0%; Pred. No. 8,4e-87;
 Matches 340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 623 AGGAAATGCAAGAGATCCAGAGCGGCATCAACAAATGGGTGTCAAATAAGACCGAAG 682
 Db 93 AGGAAATGCAAGAGATCCAGAGCGGCATCAACAAATGGGTGTCAAATAAGACCGAAG 152
 QY 683 GCCGAATCACCGATGTCTATCCCTGGAAGGCATCAATGAGTCACTGTCTGTGCTGG 742
 Db 153 GCCGAATCACCGATGTCTATCCCTGGAAGGCATCAATGAGTCACTGTCTGTGCTGG 212
 QY 743 TTACACCATTTACTTCAAGGCGCTGTGGAAGTCAAGCTTGAAGCCCTGGAACACAGGA 802
 Db 213 TTACACCATTTACTTCAAGGCGCTGTGGAAGTCAAGCTTGAAGCCCTGGAACACAGGA 272
 QY 803 AGGAATGCTTCTACAAAGCGCTGATGGAAGTGTGTCAGCATCTATGATGTACAGGAAG 862
 Db 273 AGGAATGCTTCTACAAAGCGCTGATGGAAGTGTGTCAGCATCTATGATGTACAGGAAG 332
 QY 863 GCAAGTCCGTTATCGGCGCGTGTGGAAGGCAAGGAGGCTTGAAGTTCCTTCAAG 922
 Db 333 GCAAGTCCGTTATCGGCGCGTGTGGAAGGCAAGGAGGCTTGAAGTTCCTTCAAG 392
 QY 923 GTGATGACATCACATGTGCTCTCATCTTGGCCCAAGCCTGA 962
 Db 393 GTGATGACATCACATGTGCTCTCATCTTGGCCCAAGCCTGA 432

Search completed: June 26, 2003, 22:33:45
 Job time : 356 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June '26, 2003, 22:13:10 ; Search time 214 Seconds
(without alignments)
1999.131 Million cell updates/sec

Title: US-09-828-592-6

Perfect score: 1395

Sequence: 1 atgtattccaatgtatagtag.....ccaacctgtgttaagtaa.1395

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCITUS.COMB.seq:*
6: /cgn2_6/prodata/1/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1388.6	99.5	1395	1	US-08-046-431A-1
2	1295.6	92.9	1382	3	US-09-023-173-6
3	139.8	10.0	1194	4	US-08-745-995A-28
4	139.8	10.0	1194	4	US-08-745-995A-29
5	139.8	10.0	1492	4	US-08-745-995A-25
6	139.8	10.0	1492	4	US-08-745-995A-27
7	132	9.5	1179	4	US-08-745-995A-22
8	132	9.5	1179	4	US-08-745-995A-23
9	132	9.5	1414	4	US-08-745-995A-19
10	132	9.5	1414	4	US-08-745-995A-21
11	131	9.4	1393	1	US-08-464-148-3
12	131	9.4	1393	1	US-08-385-500-3
13	131	9.4	1393	1	US-08-846-784-3
14	128.8	9.2	1197	4	US-08-745-995A-10
15	128.8	9.2	1197	4	US-08-745-995A-11
16	128.8	9.2	1358	4	US-08-745-995A-7
17	128.8	9.2	1358	4	US-08-745-995A-9
18	126.6	9.1	1191	4	US-08-745-995A-34
19	126.6	9.1	1191	4	US-08-745-995A-35
20	126.6	9.1	1454	4	US-08-745-995A-31
21	126.6	9.1	1454	4	US-08-745-995A-33
22	126.4	9.1	1191	4	US-08-745-995A-4
23	126.4	9.1	1191	4	US-08-745-995A-5
24	126.4	9.1	1584	4	US-08-745-995A-1
25	126.4	9.1	1584	4	US-08-745-995A-3
26	126.2	9.0	1260	4	US-08-745-995A-16
27	126.2	9.0	1260	4	US-08-745-995A-17

28	126.2	9.0	1838	4	US-08-745-995A-13	Sequence 13, Appl
29	126.2	9.0	1838	4	US-08-745-995A-15	Sequence 15, Appl
30	124.6	8.9	1425	1	US-08-464-148-1	Sequence 1, Appl1
31	124.6	8.9	1425	1	US-08-385-500-1	Sequence 1, Appl1
32	124.6	8.9	1425	1	US-08-846-784-1	Sequence 1, Appl1
33	119	8.5	1152	1	US-08-315-831A-13	Sequence 13, Appl
34	119	8.5	1152	1	US-08-662-318-13	Sequence 13, Appl
35	119	8.5	1152	5	PCT-US95-12509-13	Sequence 13, Appl
36	119	8.5	1316	1	US-07-755-461A-12	Sequence 12, Appl
37	119	8.5	1316	1	US-08-315-831A-12	Sequence 12, Appl
38	119	8.5	1316	1	US-08-662-318-12	Sequence 12, Appl
39	119	8.5	1316	5	PCT-US95-12509-12	Sequence 12, Appl
40	114	8.2	1366	3	US-09-200-965-1	Sequence 12, Appl
41	112	8.0	1564	3	US-08-948-997-1	Sequence 1, Appl1
42	112	8.0	1564	4	US-09-348-817A-1	Sequence 1, Appl1
43	108.4	7.8	1950	2	US-08-472-659-30	Sequence 30, Appl1
44	108.4	7.8	1950	2	US-08-474-661-30	Sequence 30, Appl1
45	108.4	7.8	1950	2	US-08-611-977-30	Sequence 30, Appl1

ALIGNMENTS

RESULT 1
US-08-046-431A-1
Sequence 1, Application US/08046431A
Patent No. 5420252
GENERAL INFORMATION:
APPLICANT: KATO, HIROYUKI
APPLICANT: YOSHITAKE, SHINJI
APPLICANT: SUZUKI, SUGURA
APPLICANT: SUZUKI, NOBORU
APPLICANT: SETO, TOSHIO
APPLICANT: NAKAO, NAOKO
APPLICANT: MIZUI, YOSHITARU
TITLE OF INVENTION: HUMAN ANTITHROMBIN III MUTANT
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLYNN, THEIL, BOUTELL & TANIS, P. C.
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/046,431A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP90488
FILING DATE: 10-APR-1992
APPLICATION NUMBER: JP1855
FILING DATE: 22-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Terryence F. Chapman
REGISTRATION NUMBER: 32549
REFERENCE/DOCKET NUMBER: Furiya Case 1286
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1395 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:

```

? ORGANISM: Homosapien
?
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 1..1395
? NAME/KEY: sig_peptide
? LOCATION: 1..96
? NAME/KEY: mat_peptide
? LOCATION: 97..1395
?
US-08-046-431A-1

```

Query Match	99.5%	Score 1388.6;	DB 1;	Length 1395;
Best Local Similarity	99.7%;	Pred. No. 0;		
Matches 1391;	Conservative	0;	Mismatches 4;	Indels 0;
				Gaps 0;

OY	1	TTGATTTCCAAATGTATAGGAACCTGAACCTCTGGAAAAAGAAAGTTATCTTTTGTC	60
Db	1	ATGATATCCAAATGTATAGGAACCTGAACCTCTGGAAAAAGAAAGTTATCTTTTGTC	60
OY	61	TTGCTGTCAATYTGAGCTTCTGGAGCTGCGTACCTGTCAACGGAGCCCTGTGACATCTGC	120
Db	61	TTGCTGTCAATYTGAGCTTCTGGAGCTGCGTACCTGTCAACGGAGCCCTGTGACATCTGC	120
OY	121	ACAGCCAAAGCCGGGACATTTCCCATGTGAATTCACATGTGATTTACCGCTCCCGGAGAA	180
Db	121	ACAGCCAAAGCCGGGACATTTCCCATGTGAATTCACATGTGATTTACCGCTCCCGGAGAA	180
OY	181	AAGGCAACTGAGATGAGGGCTCAGAAACGAAGATCCCGAGAGCCACCAACCGCGTGTTC	240
Db	181	AAGGCAACTGAGATGAGGGCTCAGAAACGAAGATCCCGAGAGCCACCAACCGCGTGTTC	240
OY	241	TGGGAACCTGTCCAAAGGCCAATTCGCGCTTGTGCTACATCTTTCTATATGACACTTGGCAGAT	300
Db	241	TGGGAACCTGTCCAAAGGCCAATTCGCGCTTGTGCTACATCTTTCTATATGACACTTGGCAGAT	300
OY	301	TCCAGAAATGACAAATGATTAACATTTTCTGTCAACCCCTAGATCTCCAGCGCTTTTGCT	360
Db	301	TCCAGAAATGACAAATGATTAACATTTTCTGTCAACCCCTAGATCTCCAGCGCTTTTGCT	360
OY	361	ATGACCAAGCTGGGTGCTGTATGACACCCCTTCAGCAACTGATGGAGGATTTTAAGTTT	420
Db	361	ATGACCAAGCTGGGTGCTGTATGACACCCCTTCAGCAACTGATGGAGGATTTTAAGTTT	420
OY	421	GACACCATCTCTGAGAAAACATCTGATCGATGTCATCTTCTTTGCGAAACGAACTGC	480
Db	421	GACACCATCTCTGAGAAAACATCTGATCGATGTCATCTTCTTTGCGAAACGAACTGC	480
OY	481	CGACTCTATCGAAAAGCCCAACAAATCTCTCAATTAATGATACGCCAATGCCCTTTTGGAA	540
Db	481	CGACTCTATCGAAAAGCCCAACAAATCTCTCAATTAATGATACGCCAATGCCCTTTTGGAA	540
OY	541	GACCAATCTCCCTTAACCTTCAATGAGACCTTACCGAGAAATCAGTGAATTTGGTATATGAGGC	600
Db	541	GACCAATCTCCCTTAACCTTCAATGAGACCTTACCGAGAAATCAGTGAATTTGGTATATGAGGC	600
OY	601	AAGCTTCAGACCCCTGTGACTTCAAGAAAAATGACAGCAATCCAGAGCGGCATTCACAACAA	660
Db	601	AAGCTTCAGACCCCTGTGACTTCAAGAAAAATGACAGCAATCCAGAGCGGCATTCACAACAA	660
OY	661	TGGGTCTCCAAATPAGACCCGAGGCGGATACCGGATGTCAATTCCTTCGGAAGCATCAAT	720
Db	661	TGGGTCTCCAAATPAGACCCGAGGCGGATACCGGATGTCAATTCCTTCGGAAGCATCAAT	720
OY	721	GAGCTCAGCTTCTGTGCTGTGCTTAAACACATTTACTTCAAGGGGCTGTGGAAGTCAAG	780
Db	721	GAGCTCAGCTTCTGTGCTGTGCTTAAACACATTTACTTCAAGGGGCTGTGGAAGTCAAG	780
OY	781	TTTCAGGCTTGAGAAACAAGAGGAAGAACTGTCTTCAAGAGCTGATGAGAGTCTGTTC	840
Db	781	TTTCAGGCTTGAGAAACAAGAGGAAGAACTGTCTTCAAGAGCTGATGAGAGTCTGTTC	840
OY	841	GCATCTATATATGACAGGAAGGCAATCCGTTATTCGGCGGTGCTGAAGCAACCCAG	900
Db	841	GCATCTATATATGACAGGAAGGCAATCCGTTATTCGGCGGTGCTGAAGCAACCCAG	900

QY	901	GTGCTTAGTGGCCCTTCAAGGATGATGATCAATCAATGATGCTCATCTTGGCCCAAGCT	960
Db	901	GTGCTTAGTGGCCCTTCAAGGATGATGATCAATCAATGATGCTCATCTTGGCCCAAGCT	960
QY	961	GAGAAAGAGCTGGCCAAAGTGGAGAAAGAACTCACTCCCAAGGTGTCTGCAGAGTGGCTG	1020
Db	961	GAGAAAGAGCTGGCCAAAGTGGAGAAAGAACTCACTCCCAAGAGTGTCTGCAGAGTGGCTG	1020
QY	1021	GATGAAATTTGAGAGAGATGATGCTGATGATGATCAATGCTCCGCTTCCGATTTGAGAGAGG	1080
Db	1021	GATGAAATTTGAGAGAGATGATGCTGATGATGATCAATGCTCCGCTTCCGATTTGAGAGAGG	1080
QY	1081	TTCAAGTTTGAAGAGAGCAGCTGCAGAGACATGAGCTGGCTTGTGATCTGTTCAGCTTGA	1140
Db	1081	TTCAAGTTTGAAGAGAGCAGCTGCAGAGACATGAGCTGGCTTGTGATCTGTTCAGCTTGA	1140
QY	1141	TTCCAAATCCCAAGGATATTGTTGCAAGAGGCCGAAATGATCACTTATGTTCTGATGATCA	1200
Db	1141	TTCCAAATCCCAAGGATATTGTTGCAAGAGGCCGAAATGATCACTTATGTTCTGATGATCA	1200
QY	1201	CATAAGCATTTCTTGAAGGTAAATGAAGAAAGCAGTAAGCAGCTGCAGATGATCCGCTGT	1260
Db	1201	CATAAGCATTTCTTGAAGGTAAATGAAGAAAGCAGTAAGCAGCTGCAGATGATCCGCTGT	1260
QY	1261	GTGATTTGTGGCCGCTTGGCTTAAACCCCAACAGGGTGACTTTCAAGGGCAACAGGCTTTT	1320
Db	1261	GTGATTTGTGGCCGCTTGGCTTAAACCCCAACAGGGTGACTTTCAAGGGCAACAGGCTTTT	1320
QY	1321	CTGGTTTTTAAAGAAAGATTCTCTGAAACAATAATTCTTCATGGCAGAGTAAAGCAAC	1380
Db	1321	CTGGTTTTTAAAGAAAGATTCTCTGAAACAATAATTCTTCATGGCAGAGTAAAGCAAC	1380
QY	1381	CCTTGTGTAAAGTAA 1395	
Db	1381	CCTTGTGTAAAGTAA 1395	

RESULT 2
US-09-023-173-6
Sequence 6, Application US/09023173
Patent No. 6066781
GENERAL INFORMATION:
APPLICANT: Sutcliffe, Thomas D.
APPLICANT: Rodriguez, Raymond L.
TITLE OF INVENTION: Production of Mature Proteins
IN PLANTS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave., Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,173
FILING DATE: 13-FEB-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/038,168
FILING DATE: 13-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Pelthory, Joanne R
REGISTRATION NUMBER: P42995
REFERENCE/DOCKET NUMBER: 0665-0007.30
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0860

TELEFAX: 650-324-0960
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1382 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 CLONE: codon opt Ramy3D-native mature A1111
 US-09-023-173-6

Query Match 92.9%; Score 1295.6; DB 3; Length 1382;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 1304; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```

QY 78 CTGGAGCTGCTGACCTGTGACGGAGCCCTGTGACATCTGACACGCCAACCCGGGGA 137
DB 57 CTGCAACAGCGGCCAGGCCCGGAGGAGCCCTGTGACATCTGACACGCCAACCCGGGGA 116
QY 138 CATTCCTCATGATCCCATGTGCAATTTACCGCTCCCGGAGAGAGAGGCACTGAGGATGA 197
DB 117 CATTCCTCATGATCCCATGTGCAATTTACCGCTCCCGGAGAGAGAGGCACTGAGGATGA 176
QY 198 GGGCTCAGAAAGAGATCCGGAGGCCCAACCGGCGTGTCTGGGAATGTCCAAAGGC 257
DB 177 GGGCTCAGAAAGAGATCCGGAGGCCCAACCGGCGTGTCTGGGAATGTCCAAAGGC 236
QY 258 CAATTCCTGCTTGTCTACCACTTTCTATAGACACCTGGAGATTCGAAATGACATTA 317
DB 237 CAATTCCTGCTTGTCTACCACTTTCTATAGACACCTGGAGATTCGAAATGACATTA 296
QY 318 TAAATTTTCTGTCAACCCCTGAGATCTCAACGCGCTTGTGATGACCAAGCTGGGTGC 377
DB 297 TAAATTTTCTGTCAACCCCTGAGATCTCAACGCGCTTGTGATGACCAAGCTGGGTGC 356
QY 378 CTGTAATGACACCTTCACAGCACTGATGAGGTATTTAATTGACACCATATCTGAGAA 437
DB 357 CTGTAATGACACCTTCACAGCACTGATGAGGTATTTAATTGACACCATATCTGAGAA 416
QY 438 AAGATCTGATGATGATCACTCTTTCTTGGCCAACTGAACTGCCGACTTATGCAAAAGC 497
DB 417 AAGATCTGATGATGATCACTCTTTCTTGGCCAACTGAACTGCCGACTTATGCAAAAGC 476
QY 498 CAAGAATCTCTCAAGTTAGTATGACGCCAATGCGCTTTTGGAGCAAAATCCCTTAACCT 557
DB 477 CAAGAATCTCTCAAGTTAGTATGACGCCAATGCGCTTTTGGAGCAAAATCCCTTAACCT 536
QY 558 CAATGAGACTTACAGAGACATGATGATGATGATGATGAGGCCAAGCTCCAGCCCTGGA 617
DB 537 CAATGAGACTTACAGAGACATGATGATGATGATGATGAGGCCAAGCTCCAGCCCTGGA 596
QY 618 CTTCAGAGAAATGACAGACATTCAGAGCGGCCCATCAAAATGAGTGTCCATTAAGAC 677
DB 597 CTTCAGAGAAATGACAGACATTCAGAGCGGCCCATCAAAATGAGTGTCCATTAAGAC 656
QY 678 CGAAGGCCGAATCACCGATGATTCCTCCGGAAGGCATCAATGAGCTCACTGTTCTGGT 737
DB 657 CGAAGGCCGAATCACCGATGATTCCTCCGGAAGGCATCAATGAGCTCACTGTTCTGGT 716
QY 738 GCTGTTTACACCATTTTACCTTCAAGGCGCTGTGAGAGTCAAAATTCAGCCCTGAGAACAC 797
DB 717 GCTGTTTACACCATTTTACCTTCAAGGCGCTGTGAGAGTCAAAATTCAGCCCTGAGAACAC 776
QY 798 AAGAGAGAACTGTTCTTACAGGCTGATGAGAGTGTGTCTGACATCTATATATGATCA 857
DB 777 AAGAGAGAACTGTTCTTACAGGCTGATGAGAGTGTGTCTGACATCTATATATGATCA 836
QY 858 GGAAGGCAAGTTCCTGTTATCGCGCGTGGCTGAAGGCAACCGAGTGTGAGTGGCTT 917
DB 837 GGAAGGCAAGTTCCTGTTATCGCGCGTGGCTGAAGGCAACCGAGTGTGAGTGGCTT 896
QY 918 CAAGGTGATGATCATCAACATGCTCTCATCTTGGCCAAAGCTGAGAAAGGCTGGCCAA 977

```

```

DB 897 CAAGGTGATGATCATCAACATGCTCTCATCTTGTGCCAAGCCTGAGAAAGCTGGCCAA 956
QY 978 GGTGAGAGAACTCACACCCAGAGTGTCTCAGAGAGTGTGATGATGATGAGAGAT 1037
DB 957 GGTGAGAGAACTCACACCCAGAGTGTCTCAGAGAGTGTGATGATGATGAGAGAT 1016
QY 1038 GATGCTGTGTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1097
DB 1017 GATGCTGTGTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1076
QY 1098 GCTGCAAGACATGAGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1157
DB 1077 GCTGCAAGACATGAGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1136
QY 1158 TGTTCAGAAAGGCCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1217
DB 1137 TGTTCAGAAAGGCCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1196
QY 1218 GGTAAATGAGAGGACATGAGAGGACATGAGAGGACATGAGAGGACATGAGAGGACAT 1277
DB 1197 GGTAAATGAGAGGACATGAGAGGACATGAGAGGACATGAGAGGACATGAGAGGACAT 1256
QY 1278 GCTAAACCCCAACAGGCTGATCTTCAAGGCCCAACAGGCTTCTGCTGTTTATTAAGA 1337
DB 1257 GCTAAACCCCAACAGGCTGATCTTCAAGGCCCAACAGGCTTCTGCTGTTTATTAAGA 1316
QY 1338 AGTTCCTGTAACATATATATCTTCAATGAGGACATGAGGACATGAGGACATGAGGACAT 1395
DB 1317 AGTTCCTGTAACATATATATCTTCAATGAGGACATGAGGACATGAGGACATGAGGACAT 1374

```

RESULT 3

US-08-745-995A-28
 Sequence 28, Application US/08745995A

Patent No. 6372887
 GENERAL INFORMATION:
 APPLICANT: Silver, Gary M.
 APPLICANT: Wisniewski, Nancy
 TITLE OF INVENTION: No. 6372887e1 Serine Protease Inhibitor
 TITLE OF INVENTION: Nucleic Acid Molecules, Proteins
 NUMBER OF INVENTIONS: 41
 NUMBER OF SEQUENCES: 41
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Heekka Corporation
 STREET: 1825 Sharp Point Drive
 CITY: Fort Collins
 STATE: Colorado
 COUNTRY: USA
 ZIP: 80525
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: Windows 95
 SOFTWARE: WordPerfect for Windows, Version 7.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/745.995A
 FILING DATE: 07-NOV-1996
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Verseer, Carol Talkington
 REGISTRATION NUMBER: 37,459
 REFERENCE/DOCKET NUMBER: PC-2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 970/493-7272
 TELEFAX: 970/484-9505
 INFORMATION FOR SEQ ID NO: 28:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1194 nucleotides
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-745-995A-28

Query Match	10.0%;	Score 139.8;	DB 4;	Length 1194;
Best Local Similarity	49.5%;	Pred. No. 2.1e-29;		
Matches 419;	Conservative	0;	Mismatches 422;	Indels 6;
				Gaps 2

OY	489	TCGAAAGCCCAACAACATCTCCCAAGTTAGTATCAGCCCAATCGCCTTTTGGAGCAAAATC	548
Db	297	TCTTAATATCACAAAAAGGTATACTCTGGAAATTGCGAATAAAGTTATGTATGGAAG	356
OY	549	CCTTACCTTCAATAGACCTTACAGACATCAGTGAAGTTGGTATATGAGCCMACTCCA	608
Db	357	CTATATCATATTAACCCACCTTCCAAAGAAAGTTGCCACCAACAAATCTTAGCTGAGCAGA	416
OY	609	GCCCCGTGACTTCAAGAAAATGACAGAGCAATCCAGAGCGGCCATCAACAAATGGAGTGC	668
Db	417	AAACTTGAACTTTGCCCAAAATGCTGAAGAGCCCTAAA---GTTATCAACACTTGGGTGA	473
OY	669	CAATTAAGACCGAAGGCCGAATCACCAGATCTATCCCTCGAAGCCATCAATAGCTAC	728
Db	474	AGAAAAAACTCATGACAAAAATTCATGATTTGTATCAAGCCGGTGATCTAGACCAAGATTTC	533
OY	729	TGTTGTGATGCGTGGTTAAACACATTTACTTCAAGAGGCCGTGGAGTCAAAAGTTACGCC	788
Db	534	AAGAAATGGTTCTTGCAATGCACTGTACTTCAAGGCTCTTGGGAAAACAATTTAGAA	593
OY	789	TGAGAAACCAAGGAAGAACTGTTCTACAGGCTATGAGAGTGTGTTCAGACTTAT	848
Db	594	GGAATACACCCAAAGCAAACTTTCTATGTTACTGAACAGAGCAAAAGAAATGTACAAAT	653
OY	849	GATGTACCAAGAGAGCAAGTTCCGTTATCGCGCGCTGCTGA--AGGCATCCAGGTGCT	905
Db	654	GATGACACATTAAAGATTAATTCGGTTATGAGAAATTTGAAGATTAAGATGCCAAGGCTGT	713
OY	906	TGAGTTGCCCTTCAAGAGTGATGATCAATCAATGGCTCATCTTGGCCAAAGCTGAGAA	965
Db	714	AGAAATGGCCCTACAGAACTCAGATTTGGCCATGTATTAATCAATTTGGCAACAGCAAAAC	773
OY	966	GAGCCTGAGCCAGGTGAGAAAGAACTCACCCCAAGGTGCTGAGAGGTGGTGATGA	1025
Db	774	TGGTCTCCCTACTCTTGAAGAAAAATTTACAAATGTGTGATTTGGCAAACTTGACTCAACG	833
OY	1026	ATTGAGAGAGATGATGCTGTGTGTCACATGCGCCGCTTCGSCATTTGAGACGGCTTAG	1085
Db	834	CATGTACTCTGTTGAAGTTATTTGGATCTGCTAAATTCAAAAATGATGCTGAATTTAA	893
OY	1086	TTTGAAGAGCAGCTGCAAGCATGSGGCTTCTGCATCTGTTCAAGCCCTGAAAAGTCCAA	1145
Db	894	TTTGAATGATCTCTGAAAAAAGTTGGGTATGTCGTATATGTTCAATGCGCTGAAAAAGCTGA	953
OY	1146	ACTCCAGATATTGTTGCAGAAAGCCGAGATACCTCTATGTCTCAGATGATTTCCATTA	1206
Db	954	TTTCAAAAGATGTTCTGAAAGATCTGATGAGATGTATATATTTTAAAGTAAATTCAAA	1013
OY	1206	GGCATTTCTTGAGATTAATGAAGAAGCCAGTAAACAGCTGCATACCGCTGTGTGAT	1265
Db	1014	AGCTTTCATTTGAAGTAAATGAAGAAGAGTGTCTAAAGCTGACAGTGCACAGGGGTATGTT	1073
OY	1266	TGCTGAGCGGTTGCTTAAACCCCAACAGGCTGACTTTCAAGGCCAACAGGCTTTCTTGCT	1325
Db	1074	AATGATGCGGTTGTATGCCAATGATGCCAATGGCTTCAATGCTGAGACATCCATCTCTGTA	1133
OY	1326	TTTTTATA 1332	
Db	1134	CTTCTTA 1140	

RESULT 4
 US-08-745-995A-29/c
 ; Sequence 29, Application US/0874595A
 ; Patent No. 6372887
 ; GENERAL INFORMATION:
 ; APPLICANT: Silver, Gary M.
 ; APPLICANT: Wisniewski, Nancy

TITLE OF INVENTION: No. 6372887el Serine Protease Inhibitor
 TITLE OF INVENTION: Nucleic Acid Molecules, Proteins
 TITLE OF INVENTION: and Uses Thereof
 NUMBER OF SEQUENCES: 41
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Heska Corporation
 STREET: 1825 Sharp Point Drive
 CITY: Fort Collins
 STATE: Colorado
 COUNTRY: USA
 ZIP: 80525
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: Windows 95
 SOFTWARE: Wordperfect for Windows, Version 7.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/
 FILING DATE: 07-NOV-1996
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Verser, Carol Talkington
 REGISTRATION NUMBER: 37,459
 REFERENCE/DOCKET NUMBER: FC-2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 970/493-7272
 TELEFAX: 970/484-9505
 INFORMATION FOR SEQ ID NO: 29:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1194 nucleotides
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA

Query Match	10.0%;	Score 139.8;	DB 4;	Length 1194;
Best Local Similarity	49.5%;	Pred. No. 2.1e-29;		
Matches 419;	Conservative 0;	Mismatches 422;	Indels 6;	Gaps 2.

OY	489	TCGAAAAGCCAAACAAATCTCCAAAGTAAGATACAGCCAAATCGCTTTTGGAGCAAAATC	548
Db	898	TCTTAAATACACAAAABAGGTGAATCTCTGGAAATTCGCAATTAAGTTATATGTATAGAAAG	839
OY	549	CTTACCTTCAATGAGACTTACAGACATCAGTGAAGTTGATATAGAGCCAAAGTCCA	608
Db	838	CTATACATTTAAACCCACCTTCAAGGAAGTTGCACCAACAAATCTTAACTGAGACAGA	779
OY	609	GCCCTCGACCTTCAAGAAAATGACAGACATCAGAGCGCCCATCAACAAATGGAGTTC	668
Db	778	AAACTTGAATCTTGGCCCAAAATGCTGAAGCGCTAAA---CTTATCAACATTTGGGTGGA	722
OY	669	CAATTAAGACCGAAGGCGGAATCAACGATGTCTATTCCTCGGAAGCCATCAATGAGCTAC	728
Db	721	AGAAAAATCATATACAAAATTCATGTGTTGATCAAAAGCGAGTATAGACACAGAAATTC	662
OY	729	TGTTCTGGGTGTGTTTAAACCATTTACTTCAAGGGCGTGGAGATCAAAAGTCAAGCCC	788
Db	661	AAGATGATGTTCTTCTCAATGCATTGTACTTCCAAAGGCTTTTGGAGAAACAAATTCMAAGA	602
OY	789	TGAGAACACAAAGNAGAACTGTTCCTAACAGGCTGATGAGAGTCTGTTCAGACATCTAT	848
Db	601	GGAAAACACCCAAAGACAAACCTTCTTATGTATTGAACACAGACAAAGAAATGTACAAAT	542
OY	849	GATGTACACAGAAAGGCAAGTTCCGTTATCGCGCGTGGCTGA---AGGACCCAGGTGCT	905
Db	541	GATCCACATTAAGATTAATTCGTTATGAGAAATTTGAAGAAATTAAGATTCGCAAGGCTGT	482
OY	906	TGAGTTGCCCTTCAAGAGTGATGATACACCATGCTCTCATCTTGGCCMAAGCCTGAGAA	965
Db	481	AGATTTGCCCTTACAGAAATCAGATTTGGGCATGTTATATCATTTTGGCAAAACAGAAAAC	422
OY	966	GAGCCTGGCCAGGTGAGAAAGAACTACCCCAAGAGTGTGCAGAGATGAGGTGATGA	1025

```

Db      421 TGGTCTCCCACTCTTGAAGAAAATTACAAAATTGATTTGGCAAACTTGACTCAAG 362
Qy      1026 ATTGAGGAGATGATGCTGTGTGTCACATGCCCGCTTCGCATTGAGAGCGGCTTCAG 1085
Db      361 CATGTACTCTGTGAAGTTATTTTGGATCTGCTTAATTAATAATTGAGTCTGAATTTAA 302
Qy      1086 TTGAGAGAGACGTGCAAGACATGGCGCTTGCATCTGTTCAGCCCTGAAAGTCCAA 1145
Db      301 TTTGATATATCTCTGAAAAGTTGGATATGTCTATATGTTCATGCTTGAAAAGCTGA 242
Qy      1146 ACTCCAGGATTTGTTGAGAGAGCCGAGATGACCTATGTCAGATGCAATTCATTA 1205
Db      241 TTTCAAGAGATTCCTTGAAGATCTGATGAGATGTTATTTCTTAAGTAAATTCAAA 182
Qy      1206 GGCATTTCTTGAAGTAAATGAAGAGCAGTGAAGCAGCTGCAAGTACCGCTGTGTAT 1265
Db      181 AGCTTTCATTAAGTAAATGAAGAGTCTGAAGCTGCAAGGCGTATGTT 122
Qy      1266 TGTGCGCGCTTGCCTAAACCCCAAGGTCATTTCAAGGCCAACAGGCTTTCCTGT 1325
Db      121 AATGATGCGTTGTATGCAATGATGCCAATGCGCTTCAATGCTGAGCATTCATCTTGA 62
Qy      1326 TTTTATA 1332
Db      61 CTTCTTA 55

```

RESULT 5

US-08-745-995A-25
Sequence 25, Application US/08745995A

GENERAL INFORMATION:

APPLICANT: Silver, Gary M.
APPLICANT: Wisniewski, Nancy
TITLE OF INVENTION: No. 6372887el Serine Protease Inhibitor
TITLE OF INVENTION: Nucleic Acid Molecules, Proteins
TITLE OF INVENTION: and Uses Thereof
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:

ADDRESSEE: Heika Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA

ZIP: 80525

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/745,995A
FILING DATE: 07-NOV-1996
CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Verwer, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: PC-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:
LENGTH: 1492 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 3..1196

US-08-745-995A-25

Query Match 10.0%; Score 139.8; DB 4; Length 1492;
Best Local Similarity 49.5%; Pred. No. 2.3e-29;
Matches 419; Conservative 0; Mismatches 422; Indels 6; Gaps 2;

```

Qy      489 TCGAAAAGCCCAAAATCTCTCAAGTATGATGAGCCATCGCCTTTTGGAGACAATC 548
Db      299 TCTTATATACAAAAGGTATGATCTGTAATTTGCCATTAAGTTATGATGAAAG 358
Qy      549 CTTTACCTTCAATGAGCTTACCAAGACATCAGTAAGTATGATGAGAACCAAGTCCA 608
Db      359 CTATACATTTAAACCCCACTTCAAGAGTTGCCACCAACAAATTTCTTACCTGAGAGA 418
Qy      609 GCCCTGAGCTTCAAGAAATGACAGACATCCAGAGCGGACCATCAACAATGCGTGTG 668
Db      419 AATCTTAACTTTGCCCAAAATGCTGAAGCGCTAAA---GTTATCAACTTTGGGTGA 475
Qy      669 CAATAGACCGAAGCCGAATCAACGATGATTCCTCTGGAAGCCATCAATGAGCTCAC 728
Db      476 AGAAAAAATCATGACAAAATTCATGATTTGATCAAGCGGTGATCTAAGACGAGATTG 535
Qy      729 TGTTCGTGCTGTGTTAAACATTTACTTCAAGGCGCTGTGAAATCAAGTTCAAGCC 788
Db      536 AAGATGTTCTTGTCAATGATTTGATCTTCAAGGCTTTTGGAGAAACATTCAGAA 595
Qy      789 TGAGAACCAAGAAAGGAACTGTTCTCAAGGCTGATGAGAGTCTGTTCAGATCTAT 848
Db      596 GAAAAACACCCAGACAAACCTTTCTATGTTACTGAACAGAACAAAGATGACGAT 655
Qy      849 GATGTACCAAGAAAGCAAGTTCCGTTATCGCGCGGTGCTGA--AGCACCCAGGTGCT 905
Db      656 GATGACATTTAAGATTAATTCGTTATGAGAAATTTGAAGATTAATGATGCAAGGCTGT 715
Qy      906 TGAATGCGCTTCAAGATGATGATCAATCATGATGCTCTTCTTGGCCAAAGCTGAA 965
Db      716 AGAATGCTCTACAGAACTCAGATTTGGCCATGTTAATCTTTTGGCAACAGCAAAAC 775
Qy      966 GAGCGTGCAGAGGTGAGAGAACTCACCCAGAGGTGTCAGAGAGTGTGATGA 1025
Db      776 TGGTCTCCCACTCTTGAAGAAAATTTCAAAATGTTGATTTGCAAAACTTGACATCAAG 835
Qy      1026 ATTGAGGAGATGATGCTGTGTGTCACATGCCCCGCTTCCGATGAGAGCGCTTCA 1085
Db      836 CATGTACTCTGTGAAGTTATTTTGGATCTGCTTAAATTTCAAAATGATGCTGAATTTAA 895
Qy      1086 TTGAGAGAGACGTGCAAGACATGGCGCTTGTGATCTGTACGCTCGTGAAGTCCAA 1145
Db      896 TTTGATATATCTCTGAAAAGTTGGATGATGCTGATATGTTCAATGCTGAAAAGCTGA 955
Qy      1146 ACTCCAGGATTTGTTGAGAGAGCCGAGATGACCTATGTCAGATGCAATTCATTA 1205
Db      956 TTTCAAGAGATTCCTTGAAGATCTGATGAGATGTTATTTCTTAAGTAAATTCAAA 1015
Qy      1206 GGCATTTCTTGAAGTAAATGAAGAGCAGTGAAGCAGCTGCAAGTACCGCTGTGTAT 1265
Db      1016 AGCTTTCATTAAGTAAATGAAGAGTCTGAAGCTGCAAGCGCGTATGTT 1075
Qy      1266 TGTGCGCGCTTGCCTAAACCCCAAGGTCATTTCAAGGCCAACAGGCTTTCCTGT 1325
Db      1076 AATGATGCGTTGTATGCAATGATGCCAATGCGCTTCAATGCTGAGCATTCATCTTGA 1135
Qy      1326 TTTTATA 1332
Db      1136 CTTCTTA 1142

```

RESULT 6

US-08-745-995A-27/c
Sequence 27, Application US/08745995A

GENERAL INFORMATION:

APPLICANT: Silver, Gary M.
APPLICANT: Wisniewski, Nancy

Matches 346; Conservative 0; Mismatches 335; Indels 3; Gaps 1;

```

Qy 652 ATCAACAATGGGTGTCATTAAGCCGAGCCGATCCAGATGTCATTCCTCGGAA 711
Db 442 ATCAACACTTGGGTGTAAGAAAACCTATACAAATTCATGATTCATTAAGCCGCT 501
Qy 712 GCCATCATGAGCTCACTGTTCTGTGCTGTTAAACCATTTTCTTCAGGGCTGTG 771
Db 502 GATCTAGACGAGATTCAAGATGTTCTTGTCAATGATGATTCATTAAGGGCTTTGG 561
Qy 772 AAGTCAAAAGTTCAGCCCTGAGAACACAGAAAGAACTGTTCTTACAGCTGATGAGAG 831
Db 562 GAGAAACATTCAGAGAGAAACACTCAAGCAAACTTTCTTATGTTACAAACAGAG 621
Qy 832 TCGTTTCAGCATCATGATGATGACAGAGAGCAAGTCCGTTATCGGGGCTGCTGA- 890
Db 622 ACAAAGATGTCAGATGATGACATTAAGATTAATTCGTTATGAGAAATTGAAAGAA 681
Qy 891 --AGGCAACCAAGCTGCTGAGTTCAGTTCCTTCAAGGTGATGACATCAGCAGTCTCATC 948
Db 682 TTAGATGCAAGGCTGATGATGATGCTTACAGAACTCAGATTTGGCATGTTATCATTT 741
Qy 949 TTGCCCCAAGCTGAGAGAGCCTGCGCAAGGTGAGAGAACTCAGCCAGAGGTCTG 1008
Db 742 TTGCCAAACAGCAAAACTGCTCTCCCGCTCTTGAAGAAAATTAACAAATGTTGACTTG 801
Qy 1009 CAGAGTGGCTGATGATGATGAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1068
Db 802 CAAACTGACTCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 861
Qy 1069 ATTGAGACGCTTCAGTTTGAAGAGAGCCTGCAAGACATGAGGCTTGTGATCTGTTTC 1128
Db 862 ATTGATCTGAATTAATTTGATGATGATCTCTGAAAAGTTGGGATCTGATATGTTT 921
Qy 1129 AGCCTGAAAAGTCCAAACTCCAGATATGTTGCAAGAGCCGAGATGACTTATGTC 1188
Db 922 GTTCTGAAAAGTCTGATTTCAAGATGCTGTAAGATCTGATGATGATGATGATGATG 981
Qy 1189 TCAGATGATTCATTAAGGCAATTTCTTGAAGTAAATGAAGAGAGAGAGAGAGCTGCA 1248
Db 982 TCTAAAGTAAATTAAGGCTTCAATTAAGTAAATGAAGAGAGAGAGAGAGCTGCACT 1041
Qy 1249 AGTACCGCTGTTGATGATGCTGCGCTTCGCTAAACCCCAAGAGGATCTTCAAGGCC 1308
Db 1042 GCCACAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1101
Qy 1309 AACAGGCTTTCCTGCTTTTATA 1332
Db 1102 GAGCATCATTCCTGACTTCTTA 1125

```

RESULT 8
US-08-745-995A-23/c
Sequence 23, Application US/08745995A
Patent No. 6372887

GENERAL INFORMATION:
APPLICANT: Silver, Gary M.

TITLE OF INVENTION: No. 6372887e1 Serine Protease Inhibitor

TITLE OF INVENTION: Nucleic Acid Molecules, Proteins

NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESS:

ADDRESSEE: Heeska Corporation

STREET: 1825 Sharp Point Drive

CITY: Fort Collins

STATE: Colorado

COUNTRY: USA

ZIP: 80525

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: Windows 95

SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/745,995A
FILING DATE: 07-NOV-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: PC-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 1179 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-745-995A-23

Query Match 9.5%; Score 132; DB 4; Length 1179;
Best Local Similarity 50.6%; Pred. No. 3.1e-27;
Matches 346; Conservative 0; Mismatches 335; Indels 3; Gaps 1;

```

Qy 652 ATCAACAATGGGTGTCATTAAGCCGAGCCGATCCAGATGTCATTCCTCGGAA 711
Db 738 ATCAACACTTGGGTGTAAGAAAACCTATACAAATTCATGATTCATTAAGCCGCT 679
Qy 712 GCCATCATGAGCTCACTGTTCTGTGCTGTTAAACCATTTTCTTCAGGGCTGTG 771
Db 678 GATCTAGACGAGATTCAAGATGTTCTTGTCAATGATGATTCATTAAGGGCTTTGG 619
Qy 772 AAGTCAAAAGTTCAGCCCTGAGAACACAGAAAGAACTGTTCTTACAGCTGATGAGAG 831
Db 618 GAGAAACATTCAGAGAGAAACACTCAAGCAAACTTTCTTATGTTACAAACAGAG 559
Qy 832 TCGTTTCAGCATCATGATGATGACAGAGAGCAAGTCCGTTATCGGGGCTGCTGA- 890
Db 558 ACAAAGATGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 499
Qy 891 --AGGCAACCAAGCTGCTGAGTTCAGTTCCTTCAAGGTGATGACATCAGCAGTCTCATC 948
Db 498 TTAGATGCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 439
Qy 949 TTGCCCCAAGCTGAGAGAGCCTGCGCAAGGTGAGAGAACTCAGCCAGAGGTCTG 1008
Db 438 TTGCCAAACAGCAAAACTGCTCTCCCGCTCTTGAAGAAAATTAACAAATGTTGACTTG 379
Qy 1009 CAGAGTGGCTGATGATGATGAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1068
Db 378 CAAACTGACTCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 319
Qy 1069 ATTGAGACGCTTCAGTTTGAAGAGAGCCTGCAAGACATGAGGCTTGTGATCTGTTTC 1128
Db 318 ATTGATCTGAATTAATTTGATGATGATCTCTGAAAAGTTGGGATCTGATATGTTT 259
Qy 1129 AGCCTGAAAAGTCCAAACTCCAGATATGTTGCAAGAGCCGAGATGACTTATGTC 1188
Db 258 GTTCTGAAAAGTCTGATTTCAAGATGCTGTAAGATCTGATGATGATGATGATGATG 199
Qy 1189 TCAGATGATTCATTAAGGCAATTTCTTGAAGTAAATGAAGAGAGAGAGAGAGCTGCA 1248
Db 198 TCTAAAGTAAATTAAGGCTTCAATTAAGTAAATGAAGAGAGAGAGAGAGCTGCACT 139
Qy 1249 AGTACCGCTGTTGATGATGCTGCGCTTCGCTAAACCCCAAGAGGATCTTCAAGGCC 1308
Db 138 GCCACAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 79
Qy 1309 AACAGGCTTTCCTGCTTTTATA 1332
Db 78 GAGCATCATTCCTGACTTCTTA 55

```


QY 1144 AACTCCAGATATTTGTCAGAGGCCAGATGACCTCTATGTCAGATGATCCAT 1203
 DB 1006 GACTTGCGGCAATGTCAGCGAGAG-----AGACTGTGTGTCCAAAGTTCGAC 1059
 QY 1204 AAGGCAATTTCTTGAGTAATGTAAGAGGCACTGAGAGCTGCAAGTCCGCTGTGG 1263
 DB 1060 AAGAGTTTGTGAGGTGATGAAAGGACCGAGGCGGCGAGCTGAGCTGCTT 1119
 QY 1264 ATTGCTGCGCGTTCCGTAAACCCCAAGGAGTACTTCAAGGCCAAGAGCCCTTCTG 1323
 DB 1120 GTAGTTGCAAGTGTCTGCAATGATGTGCCCCAGGTTCTGTCTGACACCTTTCTT 1179
 QY 1324 GTTTTATAGAGAAATCTCTGAACTATTAATTTATCTTCAAGGCGAG 1370
 DB 1180 TTCTTCATCAGGCAACAAGAGCAACAGATTCCTGTTGTGGCAG 1226

RESULT 12
 US-08-385-500-3
 ; Sequence 3, Application US/08385500
 ; Patent No. 5712117
 ; GENERAL INFORMATION:
 ; APPLICANT: Sprecher, Cindy A.
 ; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND
 ; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-3 AND CODING SEQUENCES
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend Kourie and Crew
 ; STREET: Stewart Street Tower, One Market Plaza
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: US
 ; ZIP: 94105-1493
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/385,500
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Parmelee, Steven W.
 ; REGISTRATION NUMBER: 31,590
 ; REFERENCE/DOCKET NUMBER: 13952-21
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (206) 467-9600
 ; TELEFAX: (415) 543-5043
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1393 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 112..1239
 ; OTHER INFORMATION: /product= "CYTOPLASMIC
 ; OTHER INFORMATION: ANTIPROTEINASE-3 PROTEIN"
 ; US-08-385-500-3

Query Match 9.4%; Score 131; DB 1; Length 1393;
 Best Local Similarity 49.3%; Pred. No. 6.3e-27;
 Matches 437; Conservative 0; Mismatches 435; Indels 15; Gaps 3;

QY 493 AAGCCAAATCTCTCAAGTATGATGACCAATCGCTTTTGGAGCAAAATCCCTT 552
 DB 346 AAGGCTGGACACAGATCCTGTGAGAAAGGCGCTTTTGGAGAAACTTGT 405
 QY 553 ACCTTGAATGAGCTTACGAGACATCATGAGTTGTATATAGAGCCAAAGCTCCAGCCC 612

DB 406 CAGTTCCTCTCAACCTTTAAGGAATCTGTCTTCAATTTCTACCATGAGCTGAAGAG 465
 QY 613 CTGACATTCAGAGAAATGACAGAGCATTCAGAGCGGCATCAACAATGGGTGTCCAT 672
 DB 466 CTTTCTCTTATAGAGCTGACAGAGAGTTCAGAGAAACACATCAACCTGSGTCTCMAA 525
 QY 673 AAGACCGAAGGCCGAATGACCGAGATGTCATTCCTCGAAGCCATGATGACTGCT 732
 DB 526 AAGACCGAAGGTAATTAATGAGAGTGTGTGCGGAGTACTCAATTAATGAGCAAGAAC 585
 QY 733 CTGCTGCTGTTAAACACATTTACTTCAAGGCGCTGTGAAAGTCAAGTTCAGCCCTGAG 792
 DB 586 CTGGTTCTTGTCAATGACCTTCACTTCAAGGAAAGTAAGAACCGTTTGAACAACA 645
 QY 793 AACACAGAGAAAGAACTGTTCTACAGGCTGAGAGAGTGTCTTCAAGCATCTATGATG 852
 DB 646 TACACAGAGGAATCCCTTTAAATTAACAGAGAGAGCAAAAGCCAGTGCAGATGATG 705
 QY 853 TACACAGAGGAAGGAAATCCGTTATGCGGCGCTGCTGAAG--GCACCGAGTGTGAG 909
 DB 706 TATCAGAGAGCCAGCTTAAAGCTCGCCACGCGGCGAGTGCAGCCAGCTGCTGAG 765
 QY 910 TTGCTCTTCAAGGTGATGACATCAACATGCTCTCATCTTGCCCAAGCTGAGAAAGC 969
 DB 766 CTGCTCTTCCGCGAAGAGAGTGTGCTGTGCTGTGCTGTGACGACGCGCTGAG 825
 QY 970 CTGCGCAAGGTGAGAGAAAGAACTCACCCAGAGTGTGACAGAGTG-----GCTGAT 1023
 DB 826 CTCAGACGAGTGAAGAAAGTCTTACCTTTGAGAAATTCACAGCCTGACCAAGCAAG 885
 QY 1024 GAATGAGAGAGATGATGCTGTGTGTCACATGCCCCCTTCCGATTGAGAGCGCTTC 1083
 DB 886 TGTATGAAGAGTACTGAGGTGAAAGTCTCTCTTCAAAATTTAACTACAGAGATTA 945
 QY 1084 ACTTTGAAGAGACATCCACAGACATGAGGCGCTGTGCAATGTTGAGCCCTGAAAGTCC 1143
 DB 946 GACATGGAATCTGTCTCTTCCGATTTGGAAATTTGTAGTCCCTTCAACAGGCAAGCT 1005
 QY 1144 AACTCCAGATATTTGTCAGAGGCCAGATGACCTCTATGTCAGATGATCCAT 1203
 DB 1006 GACTTGCGGCAATGTCAGCGAGAG-----AGACTGTGTGTCCAAAGTTCGAC 1059
 QY 1204 AAGGCAATTTCTTGAGTAATGTAAGAGGCACTGAGAGCTGCAAGTCCGCTGTGG 1263
 DB 1060 AAGAGTTTGTGAGGTGATGAAAGGACCGAGGCGGCGAGCTGAGCTGCTT 1119
 QY 1264 ATTGCTGCGCGTTCCGTAAACCCCAAGGAGTACTTCAAGGCCAAGAGCCCTTCTG 1323
 DB 1120 GTAGTTGCAAGTGTCTGCAATGATGTGCCCCAGGTTCTGTCTGACACCTTTCTT 1179
 QY 1324 GTTTTATAGAGAAATCTCTGAACTATTAATTTATCTTCAAGGCGAG 1370
 DB 1180 TTCTTCATCAGGCAACAAGAGCAACAGATTCCTGTTGTGGCAG 1226

RESULT 13
 US-08-846-784-3
 ; Sequence 3, Application US/08846784
 ; Patent No. 5747645
 ; GENERAL INFORMATION:
 ; APPLICANT: Sprecher, Cindy A.
 ; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND
 ; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-3 AND CODING SEQUENCES
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend Kourie and Crew
 ; STREET: Stewart Street Tower, One Market Plaza
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: US
 ; ZIP: 94105-1493
 ; COMPUTER READABLE FORM:


```

Db 460 ATCAACACTGGGTTGAGAAAAAAGCTCATGACAAAATTCATGATTGATCAAGCCGCT 519
QY 712 GCCATCAATGAGCTCACTGTTCTGCTGCTGTTAAGACCAATTTACTTCAAGGCGCTG 771
Db 520 GATCTAGACCAAGATTCAGAAATGCTTCTGTCATGATGATGATGATGATGATGATG 579
QY 772 AAGTCAAGTTCAGCCCTGAGAACACAGAGAGAACTGTTTCAAGGCTGATGAGAG 831
Db 580 GAGAAACAAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 639
QY 832 TCGTGTTCAGCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 890
Db 640 ACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 699
QY 891 --AGGACACCGAGGCTGAGTGGTCCCTTCAAGAGTATGATGATGATGATGATGAT 948
Db 700 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 759
QY 949 TTGCTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1008
Db 760 TTGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 819
QY 1009 CAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068
Db 820 CAAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 879
QY 1069 ATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1128
Db 880 ATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 939
QY 1129 AGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1188
Db 940 GTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 999
QY 1189 TCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1248
Db 1000 TCTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1059
QY 1249 AGTACCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1276
Db 1060 GCCACAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1087

```

```

RESULT 15
US-08-745-995A-11/c
; Sequence 11, Application US/08745995A
; Patent No. 6372887
; GENERAL INFORMATION:
; APPLICANT: Silver, Gary M.
; TITLE OF INVENTION: No. 6372887el Serine Protease Inhibitor
; TITLE OF INVENTION: Nucleic Acid Molecules, Proteins
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hesk Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/745,995A
; FILING DATE: 07-NOV-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459

```

```

REFERENCE/DOCKET NUMBER: FC-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1197 nucleic acid
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-745-995A-11

```

```

Query Match 9.2%; Score 128.8; DB 4; Length 1197;
Best Local Similarity 51.4%; Pred. No. 2.4e-26;
Matches 323; Conservative 0; Mismatches 302; Indels 3; Gaps 1;

```

```

QY 652 ATCAACAAATGGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 711
Db 738 ATCAACACTGGGTTGAGAAAAAAGCTCATGACAAAATTCATGATTGATCAAGCCGCT 679
QY 712 GCCATCAATGAGCTCACTGTTCTGCTGCTGTTAAGACCAATTTACTTCAAGGCGCTG 771
Db 678 GATCTAGACCAAGATTCAGAAATGCTTCTGTCATGATGATGATGATGATGATGATG 619
QY 772 AAGTCAAGTTCAGCCCTGAGAACACAGAGAGAACTGTTTCAAGGCTGATGAGAG 831
Db 618 GAGAAACAAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 559
QY 832 TCGTGTTCAGCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 890
Db 558 ACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 499
QY 891 --AGGACACCGAGGCTGAGTGGTCCCTTCAAGAGTATGATGATGATGATGATGAT 948
Db 498 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 439
QY 949 TTGCTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1008
Db 438 TTGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 379
QY 1009 CAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1068
Db 378 CAAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 319
QY 1069 ATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1128
Db 318 ATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 259
QY 1129 AGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1188
Db 258 GTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 199
QY 1189 TCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1248
Db 198 TCTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 139
QY 1249 AGTACCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1276
Db 138 GCCACAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 111

```

Search completed: June 27, 2003, 00:10:45
Job time : 217 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	1395	100.0	1395	10	US-09-8262-552-6	Sequence 6, App11
2	397.4	28.5	578	10	US-09-864-761-14838	Sequence 14838, A
3	394.2	28.3	418	9	US-09-818-3595-8157	Sequence 8157, Ap
4	390.2	28.0	395	10	US-09-864-761-11366	Sequence 11366, A
5	359.2	25.7	420	10	US-09-960-352-4934	Sequence 4934, Ap
6	341.8	24.5	409	10	US-09-960-352-9353	Sequence 9353, Ap
7	336.4	24.1	422	10	US-09-960-352-10797	Sequence 10797, A
8	334.4	24.0	418	10	US-09-960-352-12262	Sequence 12262, Ap
9	322.8	23.1	384	10	US-09-960-352-1191	Sequence 1191, Ap
10	321.8	23.1	408	10	US-09-960-352-4443	Sequence 4443, Ap
11	316	22.7	388	10	US-09-960-352-7749	Sequence 7749, Ap
12	315.4	22.6	388	10	US-09-960-352-3996	Sequence 1996, Ap
13	314.4	22.5	400	10	US-09-960-352-8899	Sequence 8899, Ap
14	311.6	22.3	368	10	US-09-960-352-14625	Sequence 14625, A
15	289.8	20.8	357	10	US-09-960-352-10061	Sequence 10061, A
16	284	20.4	433	10	US-09-960-352-8221	Sequence 8221, Ap
17	282	20.2	404	9	US-09-918-995-8263	Sequence 8263, Ap
18	281.6	20.2	328	10	US-09-960-352-3190	Sequence 3190, Ap
19	281	20.1	403	10	US-09-960-352-2960	Sequence 2960, Ap

20	276.2	19.8	361	10	US-09-960-352-2748	Sequence 2748, App
21	276.2	19.8	423	10	US-09-960-352-12011	Sequence 12011, App
22	268.8	19.3	419	10	US-09-960-352-13477	Sequence 13477, App
23	264.4	19.0	414	10	US-09-960-352-8587	Sequence 8587, App
24	264	18.9	425	10	US-09-960-352-888	Sequence 888, App
25	258	18.5	410	9	US-09-918-995-7912	Sequence 7912, App
26	254	18.2	355	10	US-09-980-107-1580	Sequence 1580, App
27	253.4	18.2	429	10	US-09-960-352-6810	Sequence 6810, App
28	253	18.1	395	9	US-09-918-995-7606	Sequence 7606, App
29	253	18.1	401	10	US-09-960-352-1087	Sequence 1087, App
30	248.2	17.8	399	10	US-09-960-352-11868	Sequence 11868, App
31	248	17.8	419	10	US-09-960-352-5400	Sequence 5400, App
32	243	17.4	416	10	US-09-960-352-14656	Sequence 14656, App
33	240.6	17.2	410	10	US-09-960-352-14810	Sequence 14810, App
34	239	17.1	387	10	US-09-960-352-10294	Sequence 10294, App
35	238.4	17.1	400	10	US-09-960-352-13253	Sequence 13253, App
36	235.2	16.9	398	10	US-09-960-352-7553	Sequence 7553, App
37	185.6	13.3	332	10	US-09-960-352-3472	Sequence 3472, App
38	178.2	12.8	267	10	US-09-980-107-2110	Sequence 2130, App
39	178.2	12.8	334	10	US-09-980-107-2294	Sequence 2294, App
40	154.2	11.1	337	10	US-09-960-352-11040	Sequence 11040, App
41	126.2	9.0	1476	9	US-09-974-298-94	Sequence 94, App
42	119	8.5	235	10	US-09-960-352-14147	Sequence 14147, App
43	119	8.5	1316	10	US-09-969-347-197	Sequence 197, App
44	119	8.5	1931	10	US-09-925-301-358	Sequence 358, App
45	118	8.5	1931	10	US-09-960-352-11778	Sequence 11778, App

ALIGNMENTS

```

RESULT 1
US-09-828-592-6
: Sequence 6, Application US/09828592
: Patent No. US20010055591A1
: GENERAL INFORMATION:
: APPLICANT: Walston, Timothy
: APPLICANT: Cooper, Scott
: APPLICANT: Revzale, Alireza
: TITLE OF INVENTION: ANTITHROMBIN H-HELIX MUTANTS
: FILE REFERENCE: 7869.10USU1
: CURRENT APPLICATION NUMBER: US/09/828,592
: CURRENT FILING DATE: 2001-04-06
: PRIOR APPLICATION NUMBER: 60/195,872
: PRIOR FILING DATE: 2000-04-07
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: SeqIdn Ver. 2.1
: SEQ ID NO 6
: LENGTH: 1395
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-828-592-6

```

Query Match	100.0%;	Score 1395;	DB 10;	Length 1395;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1395;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy ATGTAATTCAGAAAGTAGAGAACTGTAACCTCTGGAAAAAGAAAGTTTATCTTTTGTCC 60
Db 1 ATGTAATTCAGAAAGTAGAGAACTGTAACCTCTGGAAAAAGAAAGTTTATCTTTTGTCC 60
Qy TTGCTGCTCATTTGGCTTCTGGGACTGCGGTGACCTGTCAAGGGAGCCCTGTGTGACATCTGC 120
Db 61 TTGCTGCTCATTTGGCTTCTGGGACTGCGGTGACCTGTCAAGGGAGCCCTGTGTGACATCTGC 120
Qy TTGCTGCTCATTTGGCTTCTGGGACTGCGGTGACCTGTCAAGGGAGCCCTGTGTGACATCTGC 120
Db 61 TTGCTGCTCATTTGGCTTCTGGGACTGCGGTGACCTGTCAAGGGAGCCCTGTGTGACATCTGC 120
Qy ACAGCCAAAGCCCGGGACATTTCCCATGATCCCATGTGTCATTTTACCGCTCCCGGAGAG 180
Db 121 ACAGCCAAAGCCCGGGACATTTCCCATGATCCCATGTGTCATTTTACCGCTCCCGGAGAG 180
Qy AAGGCACTGAGAGTAGAGGGCTCAGAAACAGAAAGATCCCGAGGSCAACAAACCGGCTGTTC 240
Db 181 AAGGCACTGAGAGTAGAGGGCTCAGAAACAGAAAGATCCCGAGGSCAACAAACCGGCTGTTC 240

QY 241 TGGGAATGTCAGAGCCCAATTCCTGCTTGTCTACCACTTTCTATAGCACTTGGAGAT 300
 DB 241 TGGGAATGTCAGAGCCCAATTCCTGCTTGTCTACCACTTTCTATAGCACTTGGAGAT 300
 QY 301 TCCAGAAATGACAAATGATTAATTTCTGCTGACCCCTGAGTATCTCCAGGCTTTGCT 360
 DB 301 TCCAGAAATGACAAATGATTAATTTCTGCTGACCCCTGAGTATCTCCAGGCTTTGCT 360
 QY 361 ATGACCAAGCTGGGTGCTGTATGACACCTTCAGCAATCTGATGATTTAAGTT 420
 DB 361 ATGACCAAGCTGGGTGCTGTATGACACCTTCAGCAATCTGATGATTTAAGTT 420
 QY 421 GACACCATATCTGAGAAATCTGATTCAGATCTCTCTTTGCGCAACTGAACTGC 480
 DB 421 GACACCATATCTGAGAAATCTGATTCAGATCTCTCTTTGCGCAACTGAACTGC 480
 QY 481 CGACTCTATGAAAGCAAAATCTCCAGTATGATCAGCCCAATGCGCTTTTGA 540
 DB 481 CGACTCTATGAAAGCAAAATCTCCAGTATGATCAGCCCAATGCGCTTTTGA 540
 QY 541 GACAAATCCCTTACCTTCAATGAGACCTACAGACATCAGTGTGATATGAGCC 600
 DB 541 GACAAATCCCTTACCTTCAATGAGACCTACAGACATCAGTGTGATATGAGCC 600
 QY 601 AAGCTCCAGCCCTGAGCTTCAAGGAAATGACAGCAATCCAGAGCCGCTTCAACAA 660
 DB 601 AAGCTCCAGCCCTGAGCTTCAAGGAAATGACAGCAATCCAGAGCCGCTTCAACAA 660
 QY 661 TGGGTCTCAATGAGCCGAAAGCCGATCAGATGTCATTCCTCGAAGCATCAAT 720
 DB 661 TGGGTCTCAATGAGCCGAAAGCCGATCAGATGTCATTCCTCGAAGCATCAAT 720
 QY 721 GAGCTCACTGTTCTGCTGTGTATACCACTTACTTCAAGGCGCTGTGAAGTCAAG 780
 DB 721 GAGCTCACTGTTCTGCTGTGTATACCACTTACTTCAAGGCGCTGTGAAGTCAAG 780
 QY 781 TTCAAGCCCTGAGAACCAAGGAAAGAACTGTTTCAAGGCTGATGAGTGTGTTCA 840
 DB 781 TTCAAGCCCTGAGAACCAAGGAAAGAACTGTTTCAAGGCTGATGAGTGTGTTCA 840
 QY 841 GCATCATGATGATGACAGGAAAGCAATTCCTGCTGAGGCTGTGAAGCAACCG 900
 DB 841 GCATCATGATGATGACAGGAAAGCAATTCCTGCTGAGGCTGTGAAGCAACCG 900
 QY 901 GTGCTTGAAGTTCCTTCAAGGATGATGATCAGTACCATGCTTCTTCCAGGCT 960
 DB 901 GTGCTTGAAGTTCCTTCAAGGATGATGATCAGTACCATGCTTCTTCCAGGCT 960
 QY 961 GAGAAAGCTGCGCAAGGTGAGGAAAGAACTACCCCAAGGTGCTGAGAGTGTG 1020
 DB 961 GAGAAAGCTGCGCAAGGTGAGGAAAGAACTACCCCAAGGTGCTGAGAGTGTG 1020
 QY 1021 GATGATTTGAGAGAGATGATGCTGTGTGTCATGATGCTTCCGATTTGAGAGCG 1080
 DB 1021 GATGATTTGAGAGAGATGATGCTGTGTGTCATGATGCTTCCGATTTGAGAGCG 1080
 QY 1081 TTCAAGTTTGAAGAGAGCTGCAAGCAATGAGGCTTGTGATCTGTTCAAGCTGAAAG 1140
 DB 1081 TTCAAGTTTGAAGAGAGCTGCAAGCAATGAGGCTTGTGATCTGTTCAAGCTGAAAG 1140
 QY 1141 TCCAACTCCAGGATGTTGTGAGAAAGCCGAGATGACTTATGATCTCAGATGCAATTC 1200
 DB 1141 TCCAACTCCAGGATGTTGTGAGAAAGCCGAGATGACTTATGATCTCAGATGCAATTC 1200
 QY 1201 CATAGGCAATTTTGTAGGTAATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
 DB 1201 CATAGGCAATTTTGTAGGTAATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
 QY 1261 GTGATTTGCTGGCCGTTGTGCTAAACCCCAAGAGGATGCTTCAAGGCTCAAGGCTTTTC 1320
 DB 1261 GTGATTTGCTGGCCGTTGTGCTAAACCCCAAGAGGATGCTTCAAGGCTCAAGGCTTTTC 1320
 QY 1321 CTGGTTTTTATAGAGAGATTCCTCTGAGACATATATATCTTCAATGAGGAGATGAGCAAC 1380

DB 1321 CTGGTTTTTATAGAGAGATTCCTCTGAGACATATATCTTCAATGAGGAGATGAGCAAC 1380
 QY 1381 CCTTGTGTTAAGTAA 1395
 DB 1381 CCTTGTGTTAAGTAA 1395

RESULT 2
 US-09-864-761-14838
 ; Sequence 14838, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 ; FILE REFERENCE: Aecm1ca-X-1
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263,6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 14838
 ; LENGTH: 578
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AL136170.2
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 30
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 12
 ; US-09-864-761-14838

Query Match 28.5%; Score 397.4; DB 10; Length 578;
 Best Local Similarity 98.5%; Pred. NO. 1.7e-119;
 Matches 401; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 751 ATTACTTCAAGGGGCTGTGGAAGTCAAGTTGAGCCCTGAGAAACAAGAGGAAGTCTG 810
Db 125 ATTCTTTCAAGGGGCTGTGGAAGTCAAGTTGAGCCCTGAGAAACAAGAGGAAGTCTG 184
Oy 811 TTCTACAAGGCTGTATGAGAGTGTGTTTCAAGATCTATGATGTATACAGAGAGGCAAGTTC 870
Db 185 TTCTACAAGGCTGTATGAGAGTGTGTTTCAAGATCTATGATGTATACAGAGAGGCAAGTTC 244
Oy 871 CGTTATCGGCGCGTGTGCTGGAAGGCAAGGCTGTGAGTTGCTTTCAAGAGTGTATAC 930
Db 245 CGTTATCGGCGCGTGTGCTGGAAGGCAAGGCTGTGAGTTGCTTTCAAGAGTGTATAC 304
Oy 931 ATCAACCATGCTCTCATCTTGCCCAAGCTGAGAGAGCCCTGGCCCAAGGTGAGAGGA 990
Db 305 ATCAACCATGCTCTCATCTTGCCCAAGCTGAGAGAGCCCTGGCCCAAGGTGAGAGGA 364
Oy 991 CTCACCCGAGAGTGTCTCAAGAGTGTGATGATGATGAGAGAGATGATGCTGTGCTC 1050
Db 365 CTCACCCGAGAGTGTCTCAAGAGTGTGATGATGATGAGAGAGATGATGCTGTGCTC 424
Oy 1051 CACATGCCCCGCTTCCGATTTGAGAGAGCGCTTCAAGTTGAGAGAGAGCTGCAAGACTG 1110
Db 425 CACATGCCCCGCTTCCGATTTGAGAGAGCGCTTCAAGTTGAGAGAGAGCTGCAAGACTG 484
Oy 1111 GACCTTGTGATCTGTTCAGCCCTGAGAAAGTCAAACTCCAGGAT 1157
Db 485 GACCTTGTGATCTGTTCAGCCCTGAGAAAGTCAAACTCCAGGAT 531

RESULT 3

US-09-918-995-8157
; Sequence 8157, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 8157
; LENGTH: 418
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-8157

Query Match 28.3%; Score 394.2; DB 9; Length 418;
Best Local Similarity 97.9%; Pred. No. 1.5e-118;

Matches 410; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

Oy 767 CCTGGAACACAAAGGAAGTCTTCTACAAGGCTGATGAGAGTTCGTTGACACATCT 846
Db 1 CCTGGAACACAAAGGAAGTCTTCTACAAGGCTGATGAGAGTTCGTTGACACATCT 60
Oy 847 ATGATGTACAGGAAGCAAGTTCGTTATCGGCGCTGTGCTGAAGGCAAGGCTGTCT 906
Db 61 ATGATGTACAGGAAGCAAGTTCGTTATCGGCGCTGTGCTGAAGGCAAGGCTGTCT 120
Oy 907 GAGTTGCCCTTCAAGGCTGATGATCAATCAATGCTCTCATCTTTGCCCAAGCTTGAGAG 966
Db 121 GAGTTGCCCTTCAAGGCTGATGATCAATCAATGCTCTCATCTTTGCCCAAGCTTGAGAG 180
Oy 967 AGCCTGCAAGAGTGTGAGAGAGTCAAGGCTGAGAGGCTGAGAGAGTGTGATGTA 1026
Db 181 AGCCTGCAAGAGTGTGAGAGAGTCAAGGCTGAGAGGCTGAGAGAGTGTGATGTA 240
Oy 1027 TTGAGAGAGATGATCTGTGTGTCACATGCCCGCTTCCGATTGAGAGAGGCTTCAGT 1086
Db 1027 TTGAGAGAGATGATCTGTGTGTCACATGCCCGCTTCCGATTGAGAGAGGCTTCAGT 1086

Db 241 TTGAGAGAGATGATCTGTGTGTCACATGCCCGCTTCCGATTGAGAGAGGCTTCAGT 300
Oy 1087 TTGAGAGAGAGTGTGAGAGATGATGAGAGTGTGATCTGTTGACCCCTGAAGAGTCCAA 1146
Db 301 TTGAGAGAGAGTGTGAGAGATGATGAGAGTGTGATCTGTTGACCCCTGAAGAGTCCAA 359
Oy 1147 CTCACAGTATTTGTTGAGAGAGGCGGAGATGATCACTTATGCTGATGATTCATTA 1205
Db 360 CTCACAGTATTTGTTGAGAGAGGCGGAGATGATCACTTATGCTGATGATTCATTA 418

RESULT 4

US-09-864-761-31366
; Sequence 31366, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO: 31366
; LENGTH: 395
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL136170.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 30
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 12
; OTHER INFORMATION: SWISSPROT HIT: P01008, EVALUATE 4.00e-72
; OTHER INFORMATION: NT HIT: g111422599, EVALUATE 0.00e+00

OTHER INFORMATION: EST_HUMAN HIT: A1207425.1, EVALUE 0.00e+00
US-09-864-761-31366

Query Match 28.0%; Score 390.2; DB 10; Length 395;
Best Local Similarity 99.2%; Pred. No. 3.1e-117;
Matches 352; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 763 GGGCTGTGGAAGTCAAGTTCAAGCCCTGAGAACACAGAGAGAGAACTGTTCTACAAAGCT 822
DB 1 GGGCTGTGGAAGTCAAGTTCAAGCCCTGAGAACACAGAGAGAGAACTGTTCTACAAAGCT 60
QY 823 GATGAGAGATCGTGTTCAGCATCTATGATGTACAGAGAGAGAACTGTTCTACAAAGCT 882
DB 61 GATGAGAGATCGTGTTCAGCATCTATGATGTACAGAGAGAGAACTGTTCTACAAAGCT 120
QY 883 GTGGCTGAAGGACCCAGAGTGTGAGTTGCTTCAAGAGTATGACATCACTATGTC 942
DB 121 GTGGCTGAAGGACCCAGAGTGTGAGTTGCTTCAAGAGTATGACATCACTATGTC 180
QY 943 CTCATCTTCCCAAGCCCTGAGAGAGCCCTGAGAGAGAGAGAACTCAACCCAGAG 1002
DB 181 CTCATCTTCCCAAGCCCTGAGAGAGCCCTGAGAGAGAGAGAACTCAACCCAGAG 240
QY 1003 GTGCTGAGAGAGTGTGTGATGATGTGAGAGAGATGATGTGTGTCACATGCCCCG 1062
DB 241 GTGCTGAGAGAGTGTGTGATGATGTGAGAGAGATGATGTGTGTCACATGCCCCG 300
QY 1063 TTCGCGATGAGAGAGCGCTTCACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1122
DB 301 TTCGCGATGAGAGAGCGCTTCACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 1123 CTGTTGAGCCCTGAGAGAGTCCAACTCCAGGAT 1157
DB 361 CTGTTGAGCCCTGAGAGAGTCCAACTCCAGGAT 395

RESULT 5
US-09-960-352-4934
Sequence 4934, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 4934
LENGTH: 420
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 21-LIB34-080-Q1-E1-F9
US-09-960-352-4934

Query Match 25.7%; Score 359.2; DB 10; Length 420;
Best Local Similarity 91.0%; Pred. No. 4.7e-107;
Matches 352; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 200 GCTCAGAGAGAGATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 259
DB 1 GCTCAGAGAGAGATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 260 ATTCCGCTTTGTTACCATCTTTATACAGACCTGAGAGAGAGAGAGAGAGAGAGAG 319
DB 61 ATTCCGCTTTGTTACCATCTTTATACAGACCTGAGAGAGAGAGAGAGAGAGAGAG 120
QY 320 ACAATTTCTGATGAG 379
DB 121 ACAATTTCTGATGAG 180

RESULT 6
US-09-960-352-9353
Sequence 9353, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 9353
LENGTH: 409
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 40-LIB34-004-Q1-E1-B8
US-09-960-352-9353

Query Match 24.5%; Score 341.8; DB 10; Length 409;
Best Local Similarity 89.7%; Pred. No. 2.3e-101;
Matches 367; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 380 GTAATGACACCCCTCAGAGAGATGAGAGATTTAAGTTGACACCATATCTGAGAAA 439
DB 181 GTAGACACACCTCAGAGAGATGAGAGATTTAAGTTGATACCATCTCTGAGAAA 240
QY 440 CATCTGATCAGATCCATCTTCTTCTGCAAACTGAACTGCGGAGCTATATGAGAAAGCA 499
DB 241 CTTCTGATCAGATCCATCTTCTTCTGCAAACTGAACTGCGGAGCTATATGAGAAAGCA 300
QY 500 ACAATCTCCAGTTAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 559
DB 301 AATAATCTCTGAGTTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 560 ATGAGACCTTACAGAGAGATCAGTGTGATATGAGAGAGAGAGAGAGAGAGAGAGAG 619
DB 361 ATGAGACCTTACAGAGAGATCAGTGTGATATGAGAGAGAGAGAGAGAGAGAGAGAG 420

QY 456 CTTCTTCTTGGCAAGTGAAGTCCGAGCTGATGAGAGAGAGAGAGAGAGAGAGAG 515
DB 1 CTTCTTCTTGGCAAGTGAAGTCCGAGCTGATGAGAGAGAGAGAGAGAGAGAGAG 60
QY 516 AGTATCAG 575
DB 61 GGTATCAG 120
QY 576 CATCAGAGAGTGTATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 635
DB 121 CATCAGAGAGTGTATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 636 GCAATCCAG 695
DB 181 GCAATCCAG 240
QY 696 TGTCTATCTTCCGAG 755
DB 241 TGTCTATCTTCCGAG 300
QY 756 CTTCAAGAGAGCTGTGAG 815
DB 301 CTTCAAGAGAGCTGTGAG 360
QY 816 CAAGGCTGATGAG 864
DB 361 CAAGGCTGATGAG 409

```

RESULT 7
US-09-960-352-10797
; Sequence 10797, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengping
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 10797
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 46-LIB34-037-Q1-E1-D6
US-09-960-352-10797

Query Match      24.1%; Score 336.4; DB 10; Length 422;
Best Local Similarity 88.8%; Pred. No. 1.4e-99;
Matches 364; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 613 CTGACCTCAAGGAATGACAGCAATCCAGAGGCGCATCAAAATGGGTCCAT 672
DB 13 CTGGCTTCAGAGGAATGACAGCAATCCAGATTCATCAACCAATGATATCCAT 72
QY 673 AAGACCGAAGCCGCAATCAACGATGTCATTCCTCGAAGCCATGATGATCTCT 732
DB 73 AAGACTGAAGGCGCATCACTGATGTCATTCCTCCCAAGCCATCAAGATGTCATCTC 132
QY 733 CTGTGCTGCTTACACCACTTACTTCAAGGCGCTGTGGAAGTCAAGTCCCTAG 792
DB 133 CTGTGCTGCTTACACCACTTACTTCAAGGCGCTGTGGAAGTCAAGTCCCTAG 192
QY 793 AACCAAGGAAGCACTGTTCTACAAAGCTGATGAGAGTGTGTCAGCATCTATGAG 852
DB 193 AACCAAGGAAGCACTGTTCTACAAAGCTGATGAGAGTGTGTCAGCATCTATGAG 252
QY 853 TACCAAGGAAGCACTGTTCTGATGAGAGTGTGTCAGCATCTATGAGTGG 912
DB 253 TACCAAGGAAGCACTGTTCTGATGAGAGTGTGTCAGCATCTATGAGTGG 312
QY 913 CCTTCAAGGATGATGATCAATCACTGATCTCTGATCTGATCTGATGATGATG 972
DB 313 CCTTCAAGGATGATGATCAATCACTGATCTCTGATCTGATCTGATGATGATG 372
QY 973 GCCAAGGGAAGGAAGCACTCAAGGCTGAGAGTGTGCTGGA 1022
DB 373 GCCAAGGGAAGGAAGCACTCAAGGCTGAGAGTGTGCTGGA 422

```

```

; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 10-LIB34-023-Q1-E1-C5
US-09-960-352-2262

Query Match      24.0%; Score 334.4; DB 10; Length 418;
Best Local Similarity 90.3%; Pred. No. 6.3e-99;
Matches 362; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 488 ATCGAAGGCAAGCAATCCCTCCAGTAGTATGAGCAATCCCTTTGAGACCAAT 547
DB 11 ATCGAAGGCAAGCAATCCCTCCAGTAGTATGAGCAATCCCTTTGAGACCAAT 70
QY 548 CCTTCACTCAATGAGAGCTTACAGGACATCACTGATGATGATGATGATGATGATG 607
DB 71 CCTTCACTCAATGAGAGCTTACAGGACATCACTGATGATGATGATGATGATGATG 130
QY 608 AGCCCTGAGCTTCAAGGAATGCAAGCAATCCAGAGGCGCATCAAAATGGGTGT 667
DB 131 AGCCCTGAGCTTCAAGGAATGCAAGCAATCCAGAGGCGCATCAAAATGGGTGT 190
QY 668 CCAATTAAGACCGAAGGCGCATCAAGTATGATGATGATGATGATGATGATGATG 727
DB 191 CCAATTAAGACCGAAGGCGCATCAAGTATGATGATGATGATGATGATGATGATG 250
QY 728 CTGTTCGTGCTGCTGATTAACACATTTACTTCAAGGCGCTGTGGAAGTCAAGTCC 787
DB 251 CTGTTCGTGCTGCTGATTAACACATTTACTTCAAGGCGCTGTGGAAGTCAAGTCC 310
QY 788 CTGAGAACCAAGGAAGCACTGTTCTACAAAGCTGATGAGAGTGTGTCAGCATCTA 847
DB 311 CCGAAGAACCAAGGAAGCACTGTTCTACAAAGCTGATGAGAGTGTGTCAGCATCTA 370
QY 848 TGATGATCAAGGAAGCAAGTTCGATGATGATGATGATGATGATGATGATGATG 895
DB 371 TGATGATCAAGGAAGCAAGTTCGATGATGATGATGATGATGATGATGATGATG 418

RESULT 9
US-09-960-352-3191
; Sequence 3191, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengping
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 3191
; LENGTH: 384
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 14-LIB34-011-Q1-E1-D5
US-09-960-352-3191

Query Match      23.1%; Score 322.8; DB 10; Length 384;
Best Local Similarity 90.3%; Pred. No. 3.8e-95;
Matches 345; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 324 TTTCCTGTACCCCTGATGATCTCAAGGCTTTTGTGATGACCAAGCTGGTCTGTA 383
DB 2 TTTCCTGTACCCCTGATGATCTCAAGGCTTTTGTGATGACCAAGCTGGTCTGTA 61
QY 384 TGACACCTCAGGACGATGATGATGATGATGATGATGATGATGATGATGATGATG 443
DB 62 CACACACTCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 121
QY 444 TGATCAGTCACTTCTTCTTTCGCAACTGACCTGATGATGATGATGATGATGATG 503

```

Db 122 TGATGATCCACTTTTCTTGGCCAACTGAACCTGCACTCTATGAAAAACCAATTA 181
Qy 504 ATCTCCAGATAGTATGATCAGCCATGCGCTTTTGGAGACAAATCCCTTACCTTCATTA 563
Db 182 ATCTCTGAGTTGATGATCAGCCACCGTCTTTTGGAGACAAATCATTAATTCATTA 241
Qy 564 GACCTACAGAGATCAGTGGTGTATATGAGCCAAAGCTCCAGCCCTTGACTTCA 623
Db 242 GACCTACAGAGATCAGTGGTGTATATGAGCCAAAGCTCCAGCCCTTGACTTCA 301
Qy 624 GGAATATGAGAGCATTCAGAGCGGCTCAATCAATGGTGTCAATTAAGCCGAAG 683
Db 302 GGAATATGAGAGCATTCAGAGCTCAATCAATGGTGTCAATTAAGCTGAAG 361
Qy 684 CCGATACCGATGTCATTTCC 705
Db 362 GGGTATCAGTATGTCATTTCC 383

RESULT 10

US-09-960-352-4443
; Sequence 4443, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4443
; LENGTH: 408
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (330)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 19-LIB34-048-01-E1-E3
US-09-960-352-4443

Query Match 23.1%; Score 321.8; DB 10; Length 408;
Best Local Similarity 86.9%; Pred. No. 8.4e-95;
Matches 353; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy 624 GGAATATGAGAGCATTCAGAGCGGCTCAATCAATGGTGTCAATTAAGCCGAAG 683
Db 1 GGAATATGAGAGCATTCAGAGCTCAATCAATGGTGTCAATTAAGCTGAAG 60
Qy 684 CCGATACCGATGTCATTTCCGGAAGCATCAATGAGTCACTGTGTGCTGCT 743
Db 61 GCGTATCATCTGATGTCATTTCCCAAGCATCAATGAGTCACTGTGTGCTGCT 120
Qy 744 TTAACCATTTACTTCAAGGCGCTGTGAAGTCAAACTTCAAGCTTGAAGCAAGAA 803
Db 121 CAACACATTTACTTCAAGGCGCTGTGAAGTCAAACTTCAAGCTTGAAGCAAGAA 180
Qy 804 GGAAGCTGTTTACAGAGCGCTGTGAAGTCAAACTTCAAGCTTGAAGCAAGAA 863
Db 181 GGAAGCTGTTTACAGAGCGCTGTGAAGTCAAACTTCAAGCTTGAAGCAAGAA 240
Qy 864 CAAGTCCGTTATCGGCGGTGTGAAGCAAGCTTGAAGTCCGTTGAAGCTTGAAG 923
Db 241 CAAGTCCGTTATCGGCGGTGTGAAGCAAGCTTGAAGTCCGTTGAAGCTTGAAG 300
Qy 924 TGATGATCAGATGATGTCATTTCCCAAGCTTGAAGCAAGCTTGAAGCAAGTGA 983
Db 301 TGATGATCAGATGATGTCATTTCCCAAGCTTGAAGCAAGCTTGAAGCAAGTGA 360

Qy 984 GAGGAACTCAGCCCAAGAGTGTCTGAGAGTGTGCTGATGATG 1029
Db 361 ACAGGAACTCAGCCCAAGAGTGTCTGAGAGTGTGCTGATGATG 406

RESULT 11

US-09-960-352-7749
; Sequence 7749, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 7749
; LENGTH: 388
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 33-LIB34-033-Q1-E1-A2
US-09-960-352-7749

Query Match 22.7%; Score 316; DB 10; Length 388;
Best Local Similarity 88.4%; Pred. No. 6.5e-93;
Matches 343; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 282 CTATGAGCAGCTGGCAGATTCAGAGATGACATATTAATTTCTGTCACCCCTGAG 341
Db 1 CTATGAGCAGCTGGCAGATTCAGAGATGACATATTAATTTCTGTCACCCCTGAG 60
Qy 342 TATCTCCAGCGCTTTGATGAGCAAGCTGGGTCTGTATGACACCCCTCCAGACT 401
Db 61 TATCTCCAGCGCTTTGATGAGCAAGCTGGGTCTGTATGACACCCCTCCAGACT 120
Qy 402 GATGAGGATTTAAGTTGACACCATATCTGAGAAAATCTGTATGATCAGATCTTCT 461
Db 121 GATGAGGATTTAAGTTGATGATCAGATCTGAGAAAATCTGTATGATCAGATCTTCT 180
Qy 462 CTTTGCAGAACTGAAGCTGCGACTTATGAGAAAGCAACAAATCTCTCAATGTATC 521
Db 181 CTTTGCAGAACTGAAGCTGCGACTTATGAGAAAGCAACAAATCTCTCAATGTATC 240
Qy 522 AGCAATGCGCTTTTGGAGCAATCCCTTCACTTCAATGAGACTACAGACATGAG 581
Db 241 AGCAATGCGCTTTTGGAGCAATCCCTTCACTTCAATGAGACTACAGACATGAG 300
Qy 582 TGAGTTGTATATGAGCCAGCTCAGCCCTGAGCTTCAAGAAAATGACAGCAATC 641
Db 301 TGAGTTGTATATGAGCCAGCTCAGCCCTGAGCTTCAAGAAAATGACAGCAATC 360
Qy 642 CAGAGCGCATCAAAATGGGTCC 669
Db 361 TAGATTGATATCAACCAATGATATCC 388

RESULT 12

US-09-960-352-3996
; Sequence 3996, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352

[illegible]

OY	433	GAGAAAATCATGTGATCAGATCCACTCTTCTTTGGCAAACTGAACTGGCACTCATGGA	492
Db	3	GAGAAAATCTTGATCAGATCCACTCTTCTTTGGCAAACTGAACTGGCACTCATGGA	62
OY	493	AAAGCCAAATAATCCTCCCAAGTTAGTATCAGCCAAATGCGCTTTTGGAGCAAAATCCCTT	552
Db	63	AAAGCCAAATAATCCTCGAGTTGGTATCAGCCAAATCCTCTTTTGGAGCAAAATCCATT	122
OY	553	ACCTCAATGAGACCTTACAGAGATCATCATGTAGTTGGTATTTGGAGCCAAAGTCAGACC	612
Db	123	ACATTCAATGAGACCTTACAGAGATCATCATGTAGTTGGTATTTGGAGCCAAAGTCAGACC	182
OY	613	CTGGAATTCAAGAAAATGCAAGACCAATCCAGAGCGGCATCAACAAATGSGTGTCCAAAT	672
Db	183	CTGGCACTTCAAGGGAAATGCAAGAGAGTCCAGATTGATATCAACCAATGSGTATTCAAAT	242
OY	673	AAAGACCGAAGGCGGATCAACGATGTCAATTCCTTCGGAAGGCATCAATGAGTCACTGTT	732
Db	243	AAAGATGAAGGGCGATATCACTGATGTCAATTCCTCCCAACAGGCATCAATGAGTCACTGTC	302
OY	733	CTGGTGTGTTAAACACCAATTTACTTCAAGGGCTGTGGAAAGTCAAAGTTCAAGCCCTGAG	792
Db	303	CTGGTGTGTTAAACACCAATTTACTTCAAGGGCTGTGGAAAGTCAAAGTTCAAGTCCGAG	362

QY 793 AACACA 798
| | | | |
Db 363 AACACA 368

RESULT 15
US-09-960-352-10061
; Sequence 10061, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 10061
; LENGTH: 357
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 43-LIB34-023-Q1-E1-C4
US-09-960-352-10061

Query Match 20.8%; Score 289.8; DB 10; Length 357;

Best Local Similarity 88.2%; Pred. No. 2.4e-84;
Matches 315; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 519 ATCAGCCAAATCGCCCTTTTGGAGACAATCCCTTACCTTCATGAGACCTACCGAGACAT 578
| | | | |
Db 1 ATAGGCCAACCGCTTTTGGAGACAATCCCTTACCTTCATGAGACCTACCGAGACAT 60
| | | | |
QY 579 CAGTGAAGTGTATATGAGCCCAAGCTCCGCCCCCTTGACTTCAAGAAAATGCAGACA 638
| | | | |
Db 61 CAGTGAAGTGTATATGAGCCCAAGCTCCGCCCCCTTGACTTCAAGAAAATGCAGACA 120
| | | | |
QY 639 ATCCAGAGGGCCCATCAACAATGGGTGTCCATTAAGACCGAGCCGATCAGCATGT 698
| | | | |
Db 121 GTCCAGATTGACTATCAACCAATGATATCCATTAAGACTGAAGGCGTATCATGT 180
| | | | |
QY 699 CATTCCTCTGGAAGCCATCATGAGCTCACTGTCTGTGCTGTGTTAACAACATTTACTT 758
| | | | |
Db 181 CATTCCTCTGGAAGCCATCATGAGCTCACTGTCTGTGCTGTGTTAACAACATTTACTT 240
| | | | |
QY 759 CAAGGGCTCTGTGAAGTCAAGTTCAAGCCCTGAGAAACACAGAGAAGAACTGTTTACAA 818
| | | | |
Db 241 CAAGGGCTCTGTGAAGTCAAGTTCAAGCCCTGAGAAACACAGAGAAGAACTGTTTACAA 300
| | | | |
QY 819 GCGTGATGAGAGTCTGTTTCAAGCATCTATGATGTAACAGAGAAGCAAGTTCGTTA 875
| | | | |
Db 301 GCGTGATGAGAGTCTGTTTCAAGCATCTATGATGTAACAGAGAAGCAAGTTCGTTA 357
| | | | |

Search completed: June 27, 2003, 00:14:48
Job time : 232 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 16:21:51 ; Search time 1899 Seconds
(without alignments)
11897.159 Million cell updates/sec

Title: US-09-828-592-6

Perfect score: 1395

Sequence: 1 atgattccatgcatgtagt.....ccaacctgtgtcaagtaa 1395

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estmu:*
4: em_estnu:*
5: em_estcov:*
6: em_estcpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_trod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1391.8	99.8	1794	11	AF130100 Homo sapi
2	1017.6	72.9	1510	11	BC011165 Mus muscu
3	742.2	53.2	831	9	AL532135 AL532135
4	733.6	52.6	826	9	AL564291 AL564291
5	695	49.8	728	12	BG564415 BG564415
6	666	47.7	666	9	AI207425 HA2767 Hu

7	661.4	47.4	817	12	BG567696
8 <td>653.8 <td>46.9 <td>725 <td>10 <td>AV653910</td> </td></td></td></td>	653.8 <td>46.9 <td>725 <td>10 <td>AV653910</td> </td></td></td>	46.9 <td>725 <td>10 <td>AV653910</td> </td></td>	725 <td>10 <td>AV653910</td> </td>	10 <td>AV653910</td>	AV653910
9 <td>650.8 <td>46.7 <td>905 <td>9 <td>AI255692</td> </td></td></td></td>	650.8 <td>46.7 <td>905 <td>9 <td>AI255692</td> </td></td></td>	46.7 <td>905 <td>9 <td>AI255692</td> </td></td>	905 <td>9 <td>AI255692</td> </td>	9 <td>AI255692</td>	AI255692
10 <td>646 <td>46.3 <td>882 <td>13 <td>BI145342</td> </td></td></td></td>	646 <td>46.3 <td>882 <td>13 <td>BI145342</td> </td></td></td>	46.3 <td>882 <td>13 <td>BI145342</td> </td></td>	882 <td>13 <td>BI145342</td> </td>	13 <td>BI145342</td>	BI145342
11 <td>635.2 <td>45.5 <td>852 <td>12 <td>BG569095</td> </td></td></td></td>	635.2 <td>45.5 <td>852 <td>12 <td>BG569095</td> </td></td></td>	45.5 <td>852 <td>12 <td>BG569095</td> </td></td>	852 <td>12 <td>BG569095</td> </td>	12 <td>BG569095</td>	BG569095
12 <td>634.6 <td>45.5 <td>675 <td>10 <td>AV682065</td> </td></td></td></td>	634.6 <td>45.5 <td>675 <td>10 <td>AV682065</td> </td></td></td>	45.5 <td>675 <td>10 <td>AV682065</td> </td></td>	675 <td>10 <td>AV682065</td> </td>	10 <td>AV682065</td>	AV682065
13 <td>628.2 <td>45.0 <td>744 <td>12 <td>BG563372</td> </td></td></td></td>	628.2 <td>45.0 <td>744 <td>12 <td>BG563372</td> </td></td></td>	45.0 <td>744 <td>12 <td>BG563372</td> </td></td>	744 <td>12 <td>BG563372</td> </td>	12 <td>BG563372</td>	BG563372
14 <td>622.4 <td>44.6 <td>929</td> <td>9 <td>AI226504</td> </td></td></td>	622.4 <td>44.6 <td>929</td> <td>9 <td>AI226504</td> </td></td>	44.6 <td>929</td> <td>9 <td>AI226504</td> </td>	929	9 <td>AI226504</td>	AI226504
15 <td>618 <td>44.3 <td>660 <td>10 <td>AV661154</td> </td></td></td></td>	618 <td>44.3 <td>660 <td>10 <td>AV661154</td> </td></td></td>	44.3 <td>660 <td>10 <td>AV661154</td> </td></td>	660 <td>10 <td>AV661154</td> </td>	10 <td>AV661154</td>	AV661154
16 <td>613.2 <td>44.0 <td>637 <td>10 <td>AM963431</td> </td></td></td></td>	613.2 <td>44.0 <td>637 <td>10 <td>AM963431</td> </td></td></td>	44.0 <td>637 <td>10 <td>AM963431</td> </td></td>	637 <td>10 <td>AM963431</td> </td>	10 <td>AM963431</td>	AM963431
17 <td>609.4 <td>43.7 <td>613 <td>10 <td>AV647573</td> </td></td></td></td>	609.4 <td>43.7 <td>613 <td>10 <td>AV647573</td> </td></td></td>	43.7 <td>613 <td>10 <td>AV647573</td> </td></td>	613 <td>10 <td>AV647573</td> </td>	10 <td>AV647573</td>	AV647573
18 <td>605.4 <td>43.4 <td>647 <td>10 <td>AV690599</td> </td></td></td></td>	605.4 <td>43.4 <td>647 <td>10 <td>AV690599</td> </td></td></td>	43.4 <td>647 <td>10 <td>AV690599</td> </td></td>	647 <td>10 <td>AV690599</td> </td>	10 <td>AV690599</td>	AV690599
19 <td>605.2 <td>43.4 <td>825 <td>13 <td>BI219765</td> </td></td></td></td>	605.2 <td>43.4 <td>825 <td>13 <td>BI219765</td> </td></td></td>	43.4 <td>825 <td>13 <td>BI219765</td> </td></td>	825 <td>13 <td>BI219765</td> </td>	13 <td>BI219765</td>	BI219765
20 <td>596.2 <td>42.7 <td>813 <td>13 <td>BI218192</td> </td></td></td></td>	596.2 <td>42.7 <td>813 <td>13 <td>BI218192</td> </td></td></td>	42.7 <td>813 <td>13 <td>BI218192</td> </td></td>	813 <td>13 <td>BI218192</td> </td>	13 <td>BI218192</td>	BI218192
21 <td>586.8 <td>42.1 <td>824 <td>13 <td>BI148282</td> </td></td></td></td>	586.8 <td>42.1 <td>824 <td>13 <td>BI148282</td> </td></td></td>	42.1 <td>824 <td>13 <td>BI148282</td> </td></td>	824 <td>13 <td>BI148282</td> </td>	13 <td>BI148282</td>	BI148282
22 <td>583.8 <td>41.8 <td>775 <td>12 <td>BG567126</td> </td></td></td></td>	583.8 <td>41.8 <td>775 <td>12 <td>BG567126</td> </td></td></td>	41.8 <td>775 <td>12 <td>BG567126</td> </td></td>	775 <td>12 <td>BG567126</td> </td>	12 <td>BG567126</td>	BG567126
23 <td>583.6 <td>41.8 <td>660 <td>10 <td>AV691885</td> </td></td></td></td>	583.6 <td>41.8 <td>660 <td>10 <td>AV691885</td> </td></td></td>	41.8 <td>660 <td>10 <td>AV691885</td> </td></td>	660 <td>10 <td>AV691885</td> </td>	10 <td>AV691885</td>	AV691885
24 <td>580.4 <td>41.6 <td>806 <td>13 <td>BI148093</td> </td></td></td></td>	580.4 <td>41.6 <td>806 <td>13 <td>BI148093</td> </td></td></td>	41.6 <td>806 <td>13 <td>BI148093</td> </td></td>	806 <td>13 <td>BI148093</td> </td>	13 <td>BI148093</td>	BI148093
25 <td>579 <td>41.5 <td>621</td> <td>9 <td>AA706894</td> </td></td></td>	579 <td>41.5 <td>621</td> <td>9 <td>AA706894</td> </td></td>	41.5 <td>621</td> <td>9 <td>AA706894</td> </td>	621	9 <td>AA706894</td>	AA706894
26 <td>579 <td>41.5 <td>647 <td>12 <td>BG562782</td> </td></td></td></td>	579 <td>41.5 <td>647 <td>12 <td>BG562782</td> </td></td></td>	41.5 <td>647 <td>12 <td>BG562782</td> </td></td>	647 <td>12 <td>BG562782</td> </td>	12 <td>BG562782</td>	BG562782
27 <td>577.6 <td>41.4 <td>935 <td>12 <td>BF533531</td> </td></td></td></td>	577.6 <td>41.4 <td>935 <td>12 <td>BF533531</td> </td></td></td>	41.4 <td>935 <td>12 <td>BF533531</td> </td></td>	935 <td>12 <td>BF533531</td> </td>	12 <td>BF533531</td>	BF533531
28 <td>576.4 <td>41.3 <td>980 <td>13 <td>BI147005</td> </td></td></td></td>	576.4 <td>41.3 <td>980 <td>13 <td>BI147005</td> </td></td></td>	41.3 <td>980 <td>13 <td>BI147005</td> </td></td>	980 <td>13 <td>BI147005</td> </td>	13 <td>BI147005</td>	BI147005
29 <td>576 <td>41.3 <td>645 <td>10 <td>AV682911</td> </td></td></td></td>	576 <td>41.3 <td>645 <td>10 <td>AV682911</td> </td></td></td>	41.3 <td>645 <td>10 <td>AV682911</td> </td></td>	645 <td>10 <td>AV682911</td> </td>	10 <td>AV682911</td>	AV682911
30 <td>573 <td>41.1 <td>880 <td>13 <td>BI331277</td> </td></td></td></td>	573 <td>41.1 <td>880 <td>13 <td>BI331277</td> </td></td></td>	41.1 <td>880 <td>13 <td>BI331277</td> </td></td>	880 <td>13 <td>BI331277</td> </td>	13 <td>BI331277</td>	BI331277
31 <td>572.2 <td>41.1 <td>871 <td>13 <td>BI328537</td> </td></td></td></td>	572.2 <td>41.1 <td>871 <td>13 <td>BI328537</td> </td></td></td>	41.1 <td>871 <td>13 <td>BI328537</td> </td></td>	871 <td>13 <td>BI328537</td> </td>	13 <td>BI328537</td>	BI328537
32 <td>567.6 <td>40.7 <td>858 <td>13 <td>BI148112</td> </td></td></td></td>	567.6 <td>40.7 <td>858 <td>13 <td>BI148112</td> </td></td></td>	40.7 <td>858 <td>13 <td>BI148112</td> </td></td>	858 <td>13 <td>BI148112</td> </td>	13 <td>BI148112</td>	BI148112
33 <td>564.2 <td>40.4 <td>623 <td>10 <td>AV661792</td> </td></td></td></td>	564.2 <td>40.4 <td>623 <td>10 <td>AV661792</td> </td></td></td>	40.4 <td>623 <td>10 <td>AV661792</td> </td></td>	623 <td>10 <td>AV661792</td> </td>	10 <td>AV661792</td>	AV661792
34 <td>563.2 <td>40.4 <td>805</td> <td>9 <td>AI048646</td> </td></td></td>	563.2 <td>40.4 <td>805</td> <td>9 <td>AI048646</td> </td></td>	40.4 <td>805</td> <td>9 <td>AI048646</td> </td>	805	9 <td>AI048646</td>	AI048646
35 <td>563 <td>40.4 <td>805 <td>13 <td>BI246571</td> </td></td></td></td>	563 <td>40.4 <td>805 <td>13 <td>BI246571</td> </td></td></td>	40.4 <td>805 <td>13 <td>BI246571</td> </td></td>	805 <td>13 <td>BI246571</td> </td>	13 <td>BI246571</td>	BI246571
36 <td>559.6 <td>40.1 <td>867 <td>13 <td>BI218917</td> </td></td></td></td>	559.6 <td>40.1 <td>867 <td>13 <td>BI218917</td> </td></td></td>	40.1 <td>867 <td>13 <td>BI218917</td> </td></td>	867 <td>13 <td>BI218917</td> </td>	13 <td>BI218917</td>	BI218917
37 <td>559 <td>40.1 <td>727</td> <td>9 <td>AI419883</td> </td></td></td>	559 <td>40.1 <td>727</td> <td>9 <td>AI419883</td> </td></td>	40.1 <td>727</td> <td>9 <td>AI419883</td> </td>	727	9 <td>AI419883</td>	AI419883
38 <td>557.2 <td>39.9 <td>659 <td>10 <td>AV690770</td> </td></td></td></td>	557.2 <td>39.9 <td>659 <td>10 <td>AV690770</td> </td></td></td>	39.9 <td>659 <td>10 <td>AV690770</td> </td></td>	659 <td>10 <td>AV690770</td> </td>	10 <td>AV690770</td>	AV690770
39 <td>557 <td>39.9 <td>858 <td>13 <td>BI328265</td> </td></td></td></td>	557 <td>39.9 <td>858 <td>13 <td>BI328265</td> </td></td></td>	39.9 <td>858 <td>13 <td>BI328265</td> </td></td>	858 <td>13 <td>BI328265</td> </td>	13 <td>BI328265</td>	BI328265
40 <td>554.8 <td>39.8 <td>845</td> <td>9 <td>AI530232</td> </td></td></td>	554.8 <td>39.8 <td>845</td> <td>9 <td>AI530232</td> </td></td>	39.8 <td>845</td> <td>9 <td>AI530232</td> </td>	845	9 <td>AI530232</td>	AI530232
41 <td>554 <td>39.7 <td>626 <td>10 <td>AV653330</td> </td></td></td></td>	554 <td>39.7 <td>626 <td>10 <td>AV653330</td> </td></td></td>	39.7 <td>626 <td>10 <td>AV653330</td> </td></td>	626 <td>10 <td>AV653330</td> </td>	10 <td>AV653330</td>	AV653330
42 <td>552.8 <td>39.6 <td>845 <td>13 <td>BI219596</td> </td></td></td></td>	552.8 <td>39.6 <td>845 <td>13 <td>BI219596</td> </td></td></td>	39.6 <td>845 <td>13 <td>BI219596</td> </td></td>	845 <td>13 <td>BI219596</td> </td>	13 <td>BI219596</td>	BI219596
43 <td>552.2 <td>39.6 <td>771 <td>12 <td>BG617043</td> </td></td></td></td>	552.2 <td>39.6 <td>771 <td>12 <td>BG617043</td> </td></td></td>	39.6 <td>771 <td>12 <td>BG617043</td> </td></td>	771 <td>12 <td>BG617043</td> </td>	12 <td>BG617043</td>	BG617043
44 <td>551.6 <td>39.5 <td>626 <td>10 <td>AM963364</td> </td></td></td></td>	551.6 <td>39.5 <td>626 <td>10 <td>AM963364</td> </td></td></td>	39.5 <td>626 <td>10 <td>AM963364</td> </td></td>	626 <td>10 <td>AM963364</td> </td>	10 <td>AM963364</td>	AM963364
45 <td>551.6 <td>39.5 <td>872 <td>13 <td>BI328534</td> </td></td></td></td>	551.6 <td>39.5 <td>872 <td>13 <td>BI328534</td> </td></td></td>	39.5 <td>872 <td>13 <td>BI328534</td> </td></td>	872 <td>13 <td>BI328534</td> </td>	13 <td>BI328534</td>	BI328534

ALIGNMENTS

RESULT 1
LOCUS AF130100 1794 bp mRNA linear HTC 08-MAY-2001
DEFINITION Homo sapiens clone FLB1413 PRO0309 complete cds.
ACCESSION AF130100
VERSION AF130100.1 GI:11493503
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1794)
Zhang, C., Yu, Y., Zhang, S., Wei, H., Bi, J., Zhou, G., Dong, C., Zai, Y., Xu, W., Gao, F., Liu, M. and He, F.
Functional prediction of the coding sequences of 75 new genes deduced by analysis of cDNA clones from human fetal liver
Unpublished
JOURNAL 2 (bases 1 to 1794)
REFERENCE
AUTHORS Zhang, C., Yu, Y., Zhang, S., Wei, H., Bi, J., Zhou, G., Dong, C., Zai, Y., Xu, W., Gao, F., Liu, M. and He, F.
TITLE Direct Submission
JOURNAL Submitted (23-FEB-1999) Department of Experimental Hematology, Institute of Radiation Medicine, Beijing Taping Road 27, Beijing, Beijing 100850, P. R. China
FEATURES
source
1..1794
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="FLB1413"

```
repeat__region    /tissue_type="liver"  
                  /dev_stage="fetus"  
                  16. .137  
                  /rpt_family="Aiu"  
                  /rpt_type=dispersed  
CDS               234. .1628
```

BASE COUNT	540 a	429 c	420 g	405 t
ORIGIN				

Query Match	99.8%	Score 1391.8;	DB 11;	Length 1794;
Best Local Similarity	99.9%	Pred. No. 0;		
Matches 1393; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

QY	1	ATGATATTCGAATGTGATAGTGAAGCTTAACCTCTGTGAAAAAGAAAGATTATCTTTTGTC	60
Db	234	ATGATATTCGAATGTGATAGTGAAGCTTTAACTCTGTGAAAAAGAAAGTTATCTTTTGTC	293
QY	61	TTGCTGCTCATTTGGCTTCTGGGACCTGGTACCTGTCACGGAGCCCTGTGACATCTGC	120
Db	294	TTGCTGCTCATTTGGCTTCTGGGACCTGGTACCTGTCACGGAGCCCTGTGACATCTGC	353
QY	121	ACAGCCAAAGCCGCGGGACATTCCTCCATGATATCCCATGTGATTACCGCTCCCGGAGAAG	180
Db	354	ACAGCCAAAGCCGCGGGACATTCCTCCATGATATCCCATGTGATTATTCGCTCCCGGAGAAG	413
QY	181	AAGCCAACTGAGATGAGGGCTCAGAAACAGAAATCCGGAGGCCAACCGGCGTGTG	240
Db	414	AAGCCAACTGAGATGAGGGCTCAGAAACAGAAATCCGGAGGCCAACCGGCGTGTG	473
QY	241	TGGGAACCTGTCCAAAGGCCAATTCGCCGTTTGCTACACTTTTATACAGACCTGTGCAAT	300
Db	474	TGGGAACCTGTCCAAAGGCCAATTCGCCGTTTGCTACACTTTTATACAGACCTGTGCAAT	533
QY	301	TCCAAAGATGACAAATGATATACATTTTCTGTCAACCCTGATGATCTCAACGCTTTTGGT	360
Db	534	TCCAAAGATGACAAATGATATACATTTTCTGTCAACCCTGATGATCTCAACGCTTTTGGT	593
QY	361	ATGACCAAGCTGGGTGCTCTGTATACACCTCTCAGCAACTGATGAGGATTTTAAATT	420
Db	594	ATGACCAAGCTGGGTGCTCTGTATACACCTCTCAGCAACTGATGAGGATTTTAAATT	653
QY	421	GACACCATATCTGAGAAAAACATCTGATCAGATCCACTCTCTTTGGCAAACTGAACTGC	480
Db	654	GACACCATATCTGAGAAAAACATCTGATCAGATCCACTCTCTTTGGCAAACTGAACTGC	713
QY	481	CGACTCTATCGAAAAAGCCAACAAATCTCTCAAGTTAGTATCAGCCAAATCGCTTTTGGGA	540
Db	714	CGACTCTATCGAAAAAGCCAACAAATCTCTCAAGTTAGTATCAGCCAAATCGCTTTTGGGA	773
QY	541	GACAAATCTCTTAACCTTCAATGAGACTTCCAGGACATCATGTAAGTTGGTATATGAGGCC	600
Db	774	GACAAATCTCTTAACCTTCAATGAGACTTCCAGGACATCATGTAAGTTGGTATATGAGGCC	833
QY	601	AAGCTCCAGCCCTCGGACTTCAAGAAAAATGAGAGCAATCCAGAGCCGCAATCAACAAA	660
Db	834	AAGCTCCAGCCCTCGGACTTCAAGAAAAATGAGAGCAATCCAGAGCCGCAATCAACAAA	893
QY	661	TGGGTGTCGAATTAAGACCGAAAGCCGAATCACTGATGTCATTTCCCTGGAAAGCCATCAAT	720

Db	894	TTGGGTGTCATAAAGACCGAAGGCCGATATACCGATGTCAATTCCTCGGAAGCCATCAAT	953
QY	721	GAGCTCACTGTTCTGGTGTGTTTAAACAACATTTACTTTCAAGGGCCGTGTGAAGTCAAG	780
Db	954	GAGCTCACTGTTCTGGTGTGTTTAAACAACATTTACTTTCAAGGGCCGTGTGAAGTCAAG	1013
QY	781	TTGAGCCCTTGAAACAACAAGAAAGAACTGTGTTCAACAAGCTGATGTGAGAGTCTGTGTTCA	840
Db	1014	TTGAGCCCTTGAAACAACAAGAAAGAACTGTGTTCAACAAGCTGATGTGAGAGTCTGTGTTCA	1073
QY	841	GCATCTATGTATGTATCCAGAAAGCCAAAGTTCGTTATCGGCGCTGTGTAAGSACCCAG	900
Db	1074	GCATCTATGTATGTATCCAGAAAGCCAAAGTTCGTTATCGGCGCTGTGTAAGSACCCAG	1133
QY	901	GTGTTTGAATTTGCCCTTCAAGGTGATGACATCAACATGGTCCATCTTGCCCAAGCT	960
Db	1134	GTGTTTGAATTTGCCCTTCAAGGTGATGACATCAACATGGTCCATCTTGCCCAAGCT	1193
QY	961	GAGAAAGACCTGTGCGCAAGGTGAGAAAGAACTCACCCAGAGGTGTGCAAGAGTGTGCTG	1020
Db	1194	GAGAAAGACCTGTGCGCAAGGTGAGAAAGAACTCACCCAGAGGTGTGCAAGAGTGTGCTG	1253
QY	1021	GATGAATTTGAGAGAGATGATGCTGTGTGTCTCACATGCCCCGCTTCCGATTTGAGAGACGC	1080
Db	1254	GATGAATTTGAGAGAGATGATGCTGTGTGTCTCACATGCCCCGCTTCCGATTTGAGAGACGC	1313
QY	1081	TTTCAGTTTGAAGAGGACGTGCACAAGATGAGGCTTGTGCATCTGTCAGCCCTGAAAG	1140
Db	1314	TTTCAGTTTGAAGAGGACGTGCACAAGATGAGGCTTGTGCATCTGTCAGCCCTGAAAG	1373
QY	1141	TCCAAACTCCAGGATATTTGTGCGAAGAGGCGAGATGACCTTATGTCTCAGATGATTC	1200
Db	1374	TCCAAACTCCAGGATATTTGTGCGAAGAGGCGAGATGACCTTATGTCTCAGATGATTC	1433
QY	1201	CATPAGGCAATTTCTTGAAGGTAAATGAAGAAGGCAGTGAAGCACTGTGCAAGTACCGCTGT	1260
Db	1434	CATPAGGCAATTTCTTGAAGGTAAATGAAGAAGGCAGTGAAGCACTGTGCAAGTACCGCTGT	1493
QY	1261	GTGATTTGCTGGCCGTTGCTGATAAACCACAACAGGGTGACTTTCAAGGCCCAACAGGCCCTTTC	1320
Db	1494	GTGATTTGCTGGCCGTTGCTGATAAACCACAACAGGGTGACTTTCAAGGCCCAACAGGCCCTTTC	1553
QY	1321	CTGGTTTTTAAAGAGAAGTTCTCTGAACAATTTATCTTTCATGGCGAGAGTACCAAC	1380
Db	1554	CTGGTTTTTAAAGAGAAGTTCTCTGAACAATTTATCTTTCATGGCGAGAGTACCAAC	1613
QY	1381	CTTGTGCTTAAGTAA	1395
Db	1614	CTTGTGCTTAAGTAA	1628

RESULT 2	
BC011165	
LOCUS	
DEFINITION	BC011165 1510 bp mRNA linear HTC 07-AUG-2002
ACCESSION	Mus musculus. Similar to serine (or cysteine) proteinase inhibitor,
VERSION	Clade C (antithrombin), member 1, clone IMAGE:4195713, mRNA.
KEYWORDS	BC011165 BC011165.1 GI:15029882
SOURCE	HTC.
ORGANISM	house mouse.
	Mus musculus
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 1510)
AUTHORS	Straussberg R.
TITLE	Direct Submission
JOURNAL	Submitted (25-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk


```

Db      645  CATTACTTCAAGAGCCCTGTGGAAGTCAAAAGTTCAAGCCCTGAGACACAGAGAAAGAACT 586
Qy      810  GTTCTACAAAGGCTATGAGAGTCTGTTCACATCTATGATGATACCAAGAAAGCAAGTT 869
Db      565  GTTCTACAAAGGCTATGAGAGTCTGTTCACATCTATGATGATACCAAGAAAGCAAGTT 526
Qy      870  CCGTATGAGGCGGCTGAGCTGAAGGACCAAGGCTGAGTGAAGTGGCCCTCAAGAGTATGA 929
Db      525  CCGTATGAGGCGGCTGAGCTGAAGGACCAAGGCTGAGTGAAGTGGCCCTCAAGAGTATGA 466
Qy      930  CATACCAATGCTCTCATCTTTCAGCCCAAGGCTGAGAAAGGCTGAGCAAGTGAAGAGA 989
Db      465  CATACCAATGCTCTCATCTTTCAGCCCAAGGCTGAGAAAGGCTGAGCAAGTGAAGAGA 406
Qy      990  ACTACCCCAAGAGGCTGAGAGTGGCTGATGAATTTGAGAGAGATGATCTGTGTGT 1049
Db      405  ACTACCCCAAGAGGCTGAGAGTGGCTGATGAATTTGAGAGAGATGATCTGTGTGT 346
Qy      1050  CCACATGCCCCGCTCCGCTGAGAGAGAGGCTTCAAGTTGAAGAGAGCAAGCTCAAGACAT 1109
Db      345  CCACATGCCCCGCTCCGCTGAGAGAGAGGCTTCAAGTTGAAGAGAGCAAGCTCAAGACAT 286
Qy      1110  GGGGCTTGTGATCTGTTCAGCCCTGAGAAAGTCCAAAGCTCCAGGATTTGTGAGAGAG 1169
Db      285  GGGGCTTGTGATCTGTTCAGCCCTGAGAAAGTCCAAAGCTCCAGGATTTGTGAGAGAG 226
Qy      1170  CCGAATGACCTCTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1229
Db      225  CCGAATGACCTCTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 166
Qy      1230  AGGCAATGAGAGGCTGAGAGTACCGCTGTGATGATGATGATGATGATGATGATGATGATG 1289
Db      165  AGGCAATGAGAGGCTGAGAGTACCGCTGTGATGATGATGATGATGATGATGATGATGATG 106
Qy      1290  CAGGCTGATCTTCAAGGCGCAAGGCGCTTCTGCTGTTTATATAGAGAGTCTCTGAA 1349
Db      105  CAGGCTGATCTTCAAGGCGCAAGGCGCTTCTGCTGTTTATATAGAGAGTCTCTGAA 46
Qy      1350  CACTATTTATCTTC-ATGGGCAAGATGAGCAAGCTTGTGTAGT 1393
Db      45  CACTATTTATCTTCATATGAGCAAGATGAGCAAGCTTGTGTAGT 1

```

RESULT 5
 BGS64415 728 bp mRNA linear EST 10-APR-2001
 LOCUS 602584355P1 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4712245 5',
 DEFINITION mRNA sequence.
 ACCESSION BGS64415
 VERSION BGS64415.1 GI:13572067
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 728)
 NIH-MGC http://mgi.mgi.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabbs-remail.nih.gov
 Tissue Procurement: CLONETECH Laboratories, Inc.
 cDNA Library Preparation: CLONETECH Laboratories, Inc.
 cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LNCM1553 row: m column: 14
 High quality sequence stop: 726.

FEATURES
 source
 1. 728
 Location/Qualifiers

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4712245"
/lab_host="DH10B (TI phage-resistant)"
/clone_lib="NIH_MGC_76"
/notes="Organ: liver; Vector: pDNR-LIB (Clontech); Site 1:
5'fl (ggcgccgcgcgc); Site 2: 5'fl (ggcgccatagcc); 5' and
3' adaptor were used in cloning as follows: 5' adaptor
sequence: 5'-CAGCGCCATATGACC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGCGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.85
kb (range 1.0-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH MGC library."
BASE COUNT 199 a 198 c 161 g 170 t
ORIGIN

```

Query Match 49.8%; Score 695; DB 12; Length 728;
 Best Local Similarity 99.7%; Pred. No. 1.4e-186;
 Matches 717; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

```

Qy      3  GTATTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 62
Db      1  GTATTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 60
Qy      63  GCTGCTATTTGGCTTTTGGGAGTGGTGAACCTGTGACAGGAGCCCTGTGACATCTGAC 122
Db      61  GCTGCTATTTGGCTTTTGGGAGTGGTGAACCTGTGACAGGAGCCCTGTGACATCTGAC 120
Qy      123  AGCCAGGCGGCGGAGCATTTCCATGAAATCCCATGATGATTTACCGCTCCCGGAGAAAG 182
Db      121  AGCCAGGCGGCGGAGCATTTCCATGAAATCCCATGATGATTTACCGCTCCCGGAGAAAG 180
Qy      183  GGCATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 242
Db      181  GGCATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
Qy      243  GGAATCTGCAAGGCGCAATTCCTGCTTGTGATGATGATGATGATGATGATGATGATGATG 302
Db      241  GGAATCTGCAAGGCGCAATTCCTGCTTGTGATGATGATGATGATGATGATGATGATGATG 300
Qy      303  CAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 362
Db      301  CAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Qy      363  GACCAAGCTGGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 422
Db      361  GACCAAGCTGGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
Qy      423  CACCATATCTGAGAAACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 482
Db      421  CACCATATCTGAGAAACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
Qy      483  ACTCATGAGAAAGCAAAATCTCCCAAGTATGATGATGATGATGATGATGATGATGATGATG 542
Db      481  ACTCATGAGAAAGCAAAATCTCCCAAGTATGATGATGATGATGATGATGATGATGATGATG 540
Qy      543  CAATTCCTTACCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 602
Db      541  CAATTCCTTACCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Qy      603  GCTCAGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 662
Db      601  GCTCAGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
Qy      663  GATGTCATATGAGACCGAGGCGG-AATCAGCGATGATGATGATGATGATGATGATGATGATG 720
Db      660  GATGTCATATGAGACCGAGGCGGAAATCAGCGATGATGATGATGATGATGATGATGATGATG 718

```

RESULT 6
 A1207425

LOCUS AI207425 666 bp mRNA linear EST 11-NOV-1999
 DEFINITION HA2767 Human fetal liver cDNA library Homo sapiens cDNA, mRNA
 sequence.
 ACCESSION AI207425
 VERSION AI207425.1 GI:6361433
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 666)
 Yu, Y., Zhang, C., Luo, L., Ouyang, S., Li, W., Wu, J., Zhou, S., Liu, M.
 and He, F.
 Expression profile analysis of a human fetal liver cDNA library
 Unpublished (1998)
 CONTACT: Chenggang Zhang
 Beijing Institute of Radiation Medicine
 27 Taiping Road, Beijing 100850, P.R.China
 Email: zhang_chenggang@hotmail.com.
 FEATURES
 source
 1. 666
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="Human fetal liver cDNA library"
 /issue_type="Liver"
 /dev_stage="fetal"
 /lab_host="MC1061/P3"
 /note="Vector: pCDNA1"
 BASE COUNT 171 a 161 c 186 g 148 t
 ORIGIN
 Query Match 47.7%; Score 666; DB 9; Length 666;
 Best Local Similarity 100.0%; Pred. No. 2,3e-178;
 Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 647 CGGCCATCAACAAATGGTGTCCAAATTAACCGAAGCGGAATGACCATTCCT 706
 1 CGGCCATCAACAAATGGTGTCCAAATTAACCGAAGCGGAATGACCATTCCT 60
 707 CGGAACCATCAATGAGTCTACTGTTCTGGTGTGTTAAACCATTTCTCAAGGACC 766
 61 CGGAACCATCAATGAGTCTACTGTTCTGGTGTGTTAAACCATTTCTCAAGGACC 120
 767 TGTGAAGTCAAAAGTTCAAGCCCTGAGAACCAAGAAAGAACTGTTCTCAAGGCTGATG 826
 121 TGTGAAGTCAAAAGTTCAAGCCCTGAGAACCAAGAAAGAACTGTTCTCAAGGCTGATG 180
 827 GAGAGTCGTTCAGATCTATGATGTACCAAGAAAGAAAGTCCGTTATCGGCGGTGG 886
 181 GAGAGTCGTTCAGATCTATGATGTACCAAGAAAGAAAGTCCGTTATCGGCGGTGG 240
 887 CTGAAGGACCCAGAGTCTGAGTTCCTTCAAGGATGATGATGATGATGATGATGATGATG 946
 241 CTGAAGGACCCAGAGTCTGAGTTCCTTCAAGGATGATGATGATGATGATGATGATGATG 300
 947 TCTTGCCCAAGCCTGAGAAAGCCTGAGAAAGCCTGAGAAAGCCTGAGAAAGCCTGAGAAAG 1006
 301 TCTTGCCCAAGCCTGAGAAAGCCTGAGAAAGCCTGAGAAAGCCTGAGAAAGCCTGAGAAAG 360
 1007 TGCAGAGTGGCTGAGTGAATTTGAGAGAGATGATGATGATGATGATGATGATGATGATG 1066
 361 TGCAGAGTGGCTGAGTGAATTTGAGAGAGATGATGATGATGATGATGATGATGATGATG 420
 421 GCATTGAGGACGGCTTCAGTTTGAAGAGAGAGCTGCAAGACATGGGCTTGTGATCTGT 480
 1067 GCATTGAGGACGGCTTCAGTTTGAAGAGAGAGCTGCAAGACATGGGCTTGTGATCTGT 1126
 481 TCAAGCCCTGAGAAAGTCCAACTCCAGATTTGTTGAGAGAGAGAGAGAGATGATGATGATG 1186
 1127 TCAAGCCCTGAGAAAGTCCAACTCCAGATTTGTTGAGAGAGAGAGAGAGATGATGATGATG 1246
 481 TCAAGCCCTGAGAAAGTCCAACTCCAGATTTGTTGAGAGAGAGAGAGAGATGATGATGATG 540
 1187 TCAAGCCCTGAGAAAGTCCAACTCCAGATTTGTTGAGAGAGAGAGAGAGATGATGATGATG 1246

Db 541 TCTGATGATTCATTAAGGATTTCTTGAAGTAAATGAAGAGGACATGAAGACGCTG 600
 Qy 1247 CAAGTACCGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1306
 Db 601 CAAGTACCGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
 Qy 1307 CCAACA 1312
 Db 661 CCAACA 666
 RESULT 7
 BGS67696
 LOCUS 817 bp mRNA linear EST 10-APR-2001
 DEFINITION 602586477P1 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4715251 5',
 mRNA sequence.
 ACCESSION BGS67696
 VERSION BGS67696.1 GI:13575349
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 817)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: CLONETECH Laboratories, Inc.
 cDNA Library Preparation: CLONETECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L1CM1561 row: 3 column: 20
 High quality sequence stop: 808.
 Location/Qualifiers
 1. 817
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4715251"
 /clone_lib="NIH_MGC_76"
 /lab_host="DH10B (TI phage-resistant)"
 /note="Organ: liver; Vector: pNR-H1B (Clontech); Site: 1:
 Sfil (ggcgccctcgcc); Site 2: Sfil (ggcgccctcgcc); 5' and
 3' adaptors were used in cloning as follows: 5' adaptor
 sequence: 5'-CACGCCATTAAGCC-3' and 3' adaptor sequence:
 5'-ATTCTAAGCCGAGCGCGGCGGAGCATG-dt(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average inserted inserts
 kb (range 1.0-4.0 kb). 15/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA). Note: this is a NIH_MGC library."
 BASE COUNT 220 a 220 c 185 g 192 t
 ORIGIN
 Query Match 47.4%; Score 661.4; DB 12; Length 817;
 Best Local Similarity 98.7%; Pred. No. 5.5e-177;
 Matches 698; Conservative 0; Mismatches 6; Indels 3; Gaps 3;
 1 ATGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 60
 70 ATGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 129
 61 TTGCTGCTCATGCTTCTGAGAGCTGCGGACCTGTCACGGAGCCCTGTGACATCTGC 120
 130 TTGCTGCTCATGCTTCTGAGAGCTGCGGACCTGTCACGGAGCCCTGTGACATCTGC 189
 121 ACAGCCAAAGCGCGGAGATTCCTCATGATTCATGATTCATGATTCATGATTCATGATTC 180
 190 ACAGCCAAAGCGCGGAGATTCCTCATGATTCATGATTCATGATTCATGATTCATGATTC 249

181 AAGCACTGAGATGATGAGGCTCAGACGAGAGATCCGGAGGCCACCAACCGGCGTTC 240
 250 AAGGCACTGAGATGAGGCTCAGACGAGAGATCCGGAGGCCACCAACCGGCGTTC 309
 241 TGGGACTGTCAGAGGCCAATTCGCGCTTGTGTACCACTTTCTATCAGACCTGGCAGAT 300
 310 TGGGAACTGTCCAAAGCCAAATTCGCGCTTGTGTACCACTTTCTATCAGACCTGGCAGAT 369
 301 TCCAGAGATGACATGATTAACATTTTCTGTGACCCCTGAGATATCTCCACGGCTTTTGTCT 360
 370 TCCAGAGATGACATGATTAACATTTTCTGTGACCCCTGAGATATCTCCACGGCTTTTGTCT 429
 361 ATGACCAAGTGGGAGCGCTGATAGACACCTTCAGCAATGATGAGGATTTTAAGTTT 420
 430 ATGACCAAGTGGGAGCGCTGATAGACACCTTCAGCAATGATGAGGATTTTAAGTTT 489
 421 GACACCATATCTGAGAAAAACATCTGATCAGATCACTTCTTCTTGGCCAACTGAACCTGC 480
 490 GACACCATATCTGAGAAAAACATCTGATCAGATCACTTCTTCTTGGCCAACTGAACCTGC 549
 481 CGACTCTATCGAAAAAGCCAAATCTTCAGATGATCAGCAATCGCCTTTTGTGA 540
 550 CGACTCTATCGAAAAAGCCAAATCTTCAGATGATCAGCAATCGCCTTTTGTGA 609
 541 GACAAATCCCTTACCTTCAATGAGACTACGAGACATGAGTGGTATATGAGGCC 600
 610 GACAAATCCCTTACCTTCAATGAGACTACGAGACATGAGTGGTATATGAGGCC 669
 601 AAGCTCCAGCCCTT-GAGACTTCAGAAAAA--TGACAGAGCAATCGAGCGCCATCAACA 658
 670 AAGCTCCAGCCCTTGGAGCTTCAGAAAAAATCGAGAGCAATCGAGCGCCATCAACA 729
 659 AATGGGTGTCATAG-ACCGAAGCGGCAATCACCAGTGTCTTTC 704
 730 AATGGGTGTCATAGAACCGAAGCGGCAATCACCAGTGTCTTTC 776

RESULT 8
 AV653910 725 bp mRNA linear EST 15-JAN-2002
 LOCUS AV653910 GLC Homo sapiens cDNA clone GLCDPH09 3', mRNA sequence.
 DEFINITION AV653910
 ACCESSION AV653910
 VERSION AV653910.1 GI:9874924
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
 AUTHORS Xu X., Huang J., Xu Z., Qian B., Zhu Z., Yan Q., Cai T., Zhang X., Xiao H., Qu J., Liu F., Huang Q., Cheng Z., Li N., Du J., Hu W., Shen K., Lu G., Fu G., Zhong M., Xu S., Gu W., Huang W., Zhao X., Hu G., Gu J., Chen Z., and Han Z.
 TITLE Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
 MEDLINE 21625106
 COMMENT Contact: Zeguang Han
 Chinese National Human Genome Center at Shanghai
 351 Guoshoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919(ex. 45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn
 This clone is available at CHGC in Shanghai.

FEATURES

source
 1..725
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="GLCDPH09"
 /clone_1ib="GLC"

/tissue type="corresponding non cancerous liver tissue"
 /dev stage="Adult"
 /lab_host="SOLR"
 /note="Vector: pluescript sk(-), Site_1: EcoRI; Site_2: XhoI"

Query Match 46.9%; Score 653.8; DB 10; Length 725;
 Best Local Similarity 98.6%; Pred. No. 7.4e-175;
 Matches 680; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

BASE COUNT 195 a 177 c 191 g 160 t 2 others
 ORIGIN
 459 CTTCTTGGCCAACTGAACTCCGACTCTATCGAAAAAGCCAAATCTCCAGTATG 518
 1 CTTCTTGGCCAACTGAACTCCGACTCTATCGAAAAAGCCAAATCTCCAGTATG 60
 519 ATCAGCCAAATGCGCTTTTGGAGCAAAATCCCTTACCTTCAATGAGACTACAGACAT 578
 61 ATCAGCCAAATGCGCTTTTGGAGCAAAATCCCTTACCTTCAATGAGACTACAGACAT 120
 579 CAGTGAAGTGTATATGAGCCAAAGCTTCAGCCCTTGAATCTTCAAGAAAAATCGAGACA 638
 121 CAGTGAAGTGTATATGAGCCAAAGCTTCAGCCCTTGAATCTTCAAGAAAAATCGAGACA 180
 639 ATCCAGAGCGGCCATCAACAAATGGGTGTCATTAAGACGAGCGGCAATCACCGATGT 698
 181 ATCCAGAGCGGCCATCAACAAATGGGTGTCATTAAGACGAGCGGCAATCACCGATGT 240
 699 CATTCCCTCGGAGACCATCAATGAGCTCACTGTTCTGTGTGTTTAAACACATTACTT 758
 241 CATTCCCTCGGAGACCATCAATGAGCTCACTGTTCTGTGTGTTTAAACACATTACTT 300
 759 CAAGGCGCTGTGAAAGTCAAAAGTTCAAGCTTCGAGAAACAAAGAAAGAACTTTCTACA 818
 301 CAAGGCGCTGTGAAAGTCAAAAGTTCAAGCTTCGAGAAACAAAGAAAGAACTTTCTACA 360
 819 GAGTGAAGAGAGTGTGTTGAGCATCTATGATGATGACAGAGAGCAAGTCCGTTATCG 878
 361 GAGTGAAGAGAGTGTGTTGAGCATCTATGATGATGACAGAGAGCAAGTCCGTTATCG 420
 879 GCGGCTGCTGAGAGCAACCCAGGTGCTTGAATGCTCCCTTCAAGAGTGAATGACATCCAT 938
 421 GCGGCTGCTGAGAGCAACCCAGGTGCTTGAATGCTCCCTTCAAGAGTGAATGACATCCAT 480
 939 GGTCTCTATCTTGGCCAAAGCTTGAAGAGAGCTTGGCCCAAGTGAAGAACTTCAACCC 998
 481 GGTCTCTATCTTGGCCAAAGCTTGAAGAGAGCTTGGCCCAAGTGAAGAACTTCAACCC 540
 999 AGAGGTGCTGAGAGAGTGTGATGATGATGAGAGAGTGTGCT-GGTGTCCACATGC 1057
 541 AGAGGTGCTGAGAGAGTGTGATGATGATGAGAGAGTGTGCTGGTGTCCACATGC 600
 1058 CCCGCTTCCGATGAGAGCGCTTCAAGTTGAAGAGAGCACTGCACATGCG-GCCTT 1116
 601 CCCGCTTCCGATGAGAGCGCTTCAAGTTGAAGAGAGCACTGCACATGCGGCTT 660
 1117 GTCGATCTGTTCACGCTTGAAGAGTCCAA 1146
 661 GTCGATCTGTTCACGCTTGAAGAGTCCAA 690

RESULT 9
 A1255692 905 bp mRNA linear EST 12-NOV-1998
 LOCUS A1255692
 DEFINITION u156e10.y1 Sugano mouse liver m1a Mus musculus cDNA clone
 IMAGE:1886442 5' similar to gb:U68793_fna1 ANTIHROMBIN-III
 PRECURSOR (HUMAN);, mRNA sequence.
 ACCESSION A1255692
 VERSION A1255692.1 GI:3863217
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

QY 301 TCACAAATGACATGATTAACATTTTCTGTCAACCCCTGAGTATCTCCAGGCTTTTGGT 360
 DB 349 TCACAAATGACATGATTAACATTTTCTGTCAACCCCTGAGTATCTCCAGGCTTTTGGT 408
 QY 361 ATGACCAAGCTGGGCTGTATGACACCTTCAGCA--CTATGGAGGATTTAAGTT 419
 DB 409 ATACCAAGCTGGGCTGTATGACACCTTCAGCAACCTTATGGAGGATTTAAGTT 468
 QY 420 TGACACCATATCTGAGAAAAATCTGATGATGATCACTTCTTTGCCAACTGAATG 479
 DB 469 TGACACCATATCTGAGAAAAATCTGATGATGATCACTTCTTTGCCAACTGAATG 528
 QY 480 CCGACTCTATCGAAAGCCCAAAATCTCC--AAGTTAGTATGAGCAATCGCTTTTGG 538
 DB 529 CCGACTCTATCGAAAGCCCAAAATCTCCAAAGTTAGTATGAGCAATCGCTTTTGG 588
 QY 539 GAGACAAATCCCTTACC--TTCAATGAGACCTACAGGACATGAGTGGTATATGGA 597
 DB 589 GAGACAAATCCCTTACCCTTCAATGAGACCTACAGGACATGAGTGGTATATGGA 648
 QY 598 GCCAAGCTCCAGCCCTCGAATTCA--GAAAAATGACAGCAATCCAGAGCGGCATC 654
 DB 649 GCCAAGCTCCAGCCCTCGAATTCAAGGCAATGAGCAATCCAGAGCGGCATC 708
 QY 655 AACAAATGGGTCTCAATAGA--CCGAGGCGCAATCCAGATGATTTCCCTCGAAGC 713
 DB 709 AACAAATGGGTCTCAATAGACCCGAGGCGCAATCCAGATGATTTCCCTCGAAGC 768
 QY 714 CATCATGAGCTCACTGTTCTGGTGCTGTTAACACCATTTACTTCAAGGCGCTTGGA 773
 DB 769 CATC-ATGAGCTACTG-CTGGTGCTGAT-ACACCGAGTACTTCAAGG--CTGGA 823
 QY 774 GTCAAGTTGAGCC 787
 DB 824 GTCAAGTTGAGCC 837

RESULT 12
 LOCUS AV682065 675 bp mRNA linear EST 16-JUN-2002
 DEFINITION AV682065 GRX Homo sapiens cDNA clone GKBAB07 5', mRNA sequence.
 ACCESSION AV682065
 VERSION AV682065.1 GI:10283928
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 675)

TITLE
 Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X., Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W., Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X., Hu,G., Gu,J., Chen,Z. and Han,Z.
 Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver

JOURNAL
 MEDLINE Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
 COMMENT 21625106
 Contact: Zeguang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919 (ex.45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn

FEATURES
 source
 Location/Qualifiers
 1..675
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="GKBAB07"
 /clone_1b="GKB"
 /tissue_type="hepatocellular carcinoma"

BASE COUNT 180 a 167 c 177 g 147 t 4 others
 ORIGIN
 Query Match 45.5%; Score 634.6; DB 10; Length 675;
 Best Local Similarity 97.8%; Pred. No. 2.1e-169;
 Matches 662; Conservative 0; Mismatches 13; Indels 2; Gaps 2;
 /dev_stage="Adult"
 /lab_host="SOLR"
 /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"

QY 459 CTTCTTGGCAACGAACTGGCGACTATGAGAAAGCAACATCTCCAAAGTTAGT 518
 DB 1 CTTCTTGGCAACGAACTGGCGACTATGAGAAAGCAACATCTCCAAAGTTAGT 60
 QY 519 ATACGCAATCGCCTTTTGGAGACAAATCCCTTCACTTCAATGAGACTTACAGACAT 578
 DB 61 ATACGCAATCGCCTTTTGGAGACAAATCCCTTCACTTCAATGAGACTTACAGACAT 120
 QY 579 CAGTGAATTGTATATGAGCCAGCTCCAGCTCCGACTTCAAGAAATGCAAGACA 638
 DB 121 CAGTGAATTGTATATGAGCCAGCTCCAGCTCCGACTTCAAGAAATGCAAGACA 180
 QY 639 ATCCAGAGCGGCATCAACAAATGGGTGTCATTAAGACCGAAGCCGATCAGCATGT 698
 DB 181 ATCCAGAGCGGCATCAACAAATGGGTGTCATTAAGACCGAAGCCGATCAGCATGT 240
 QY 699 CATTCCTCGAAGCATCAATGAGCTCACTGTTCTGGTGCTGTTAAACCATTTACTT 758
 DB 241 CATTCCTCGAAGCATCAATGAGCTCACTGTTCTGGTGCTGTTAAACCATTTACTT 300
 QY 759 CAAGGCGCTGTGGAAGTCAAGTTCAAGCTTGAAGACACAGAAAGAACTGTTTCA 818
 DB 301 CAAGGCGCTGTGGAAGTCAAGTTCAAGCTTGAAGACACAGAAAGAACTGTTTCA 360
 QY 819 GGCTGATGAGAGTGTGTGTCAGCATATGATATGACAGAAAGCAAGTTCCGTTATG 878
 DB 361 GGCTGATGAGAGTGTGTGTCAGCATATGATATGACAGAAAGCAAGTTCCGTTATG 420
 QY 879 GCGCGTGTGTAAGGACCCAGAGTCTTGAAGTTCCTTCAAGGTGATGATCATCAT 938
 DB 421 GCGCGTGTGTAAGGACCCAGAGTCTTGAAGTTCCTTCAAGGTGATGATCATCAT 480
 QY 939 GGTCTCATCTTGGCCCAAGCTTGAAGAGAGCTTGGCCAAAGTGAAGAACTCAACCC 998
 DB 481 GGTCTCATCTTGGCCCAAGCTTGAAGAGAGCTTGGCCAAAGTGAAGAACTCAACCC 540
 QY 999 AAGAGTGTGACAGAGTGGCTGATGAATGAGAGAGATGATGCTGATGATGATG 1058
 DB 541 AAGAGTGTGACAGAGTGGCTGATGAATGAGAGAGATGATGCTGATGATGATG 600
 QY 1059 CCGCTTCCGATTAAGAGACGCGCTTGAAGTGAAGAGACGCTCAAGCATGAGGCTTTGT 1118
 DB 601 CCG-TTTGGCATTAAGAGACGCGCTTGAAGTGAAGAGACGCTCAAGCATGAGG-CTTGT 658
 QY 1119 CGATCTGTTCAAGCCTTG 1135
 DB 659 CGATCTGTTCAAGCCTTG 675

RESULT 13
 LOCUS BG563372 744 bp mRNA linear EST 10-APR-2001
 DEFINITION BG563372 602582428F1 NIH_MGC_76 Homo sapiens cDNA clone IMAGE:4710199 5', mRNA sequence.
 ACCESSION BG563372
 VERSION BG563372.1 GI:13571024
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 744)

Query Match 44.6%; Score 622.4; DB 9; Length 929;
 Best Local Similarity 81.3%; Pred. No. 7.8e-166;
 Matches 758; Conservative 0; Mismatches 171; Indels 3; Gaps 3;

442 TCGATCAGATCAGCTCTTCTTCTTGGCAAACTGACGCGCATCTATCGAAAAGCCAC 501
 1 TCGACAGATCAGCTCTTCTTCTTGGCAAACTGACGCGCATCTATCGAAAAGCCAC 60
 502 AATCTCCAGAGTATGATCAGCAATCGCTTTTGGAGCAAAATCCCTTACCTTCAAT 561
 61 AAGTCTCTGACTTGGATCAGCAACCGCTTTTGGAGCAAAATCCCTTACCTTCAAC 120
 562 GAGACTTACAGACATCAGTGTGTGTATGAGCCAGCTCCAGCCCTTGAGCTTC 621
 121 GAGAGTATCAGATGT 180
 622 AAGCAATATGACAGAGATCAGAGCGGCGCATCAAAATGGGTGTCAATAAGACCGAA 681
 181 AAGCAATATCAGAGATCAGAGCGGCGCATCAAAATGGGTGTCAATAAGACCGAA 240
 682 GGGCGATCAGAGATCAGAGCGGCGCATCAAAATGGGTGTCAATAAGACCGAA 741
 241 GGGCGATCAGAGATCAGAGCGGCGCATCAAAATGGGTGTCAATAAGACCGAA 300
 742 GTTAAACACCATTTACTTCAAGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 801
 301 GTTAAACACCATTTACTTCAAGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
 802 AAGCAATGT 861
 361 AAGCAATGT 420
 862 GGGCAATGT 921
 421 GGGCAATGT 480
 922 GGGCAATGT 981
 481 GGGCAATGT 538
 982 GAGCAAGACTCAGCCCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1041
 539 GAGCAAGACTCAGCCCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 598
 1042 CTGT 1101
 599 CTGT 658
 1102 CAGCAATGT 1161
 659 CAGCAATGT 717
 1162 GCAGAGGCGGAGATGACCTCTATGTCTCAGATGATGATTCATTAAGCATTTCTTGA 1221
 718 GCTGTATGCAAGAGACACCTCTATGTCTCAGATGATGATTCATTAAGCATTTCTTGA 777
 1222 AATGAAGAGGAG 1281
 778 AATGAAGAGGAG 837
 1282 AATCCCAAG 1341
 838 AATCCCAAG 897
 1342 CTTCTGAACACTATTTATCTTCAATGAGAGAGT 1373
 898 TGCAGTGCAGCTATTTACTTCAATGAGAGAGT 929

RESULT 15
 AV661154 660 bp mRNA linear EST 16-JAN-2002
 LOCUS AV661154 GIC Homo sapiens cDNA clone GICPC05 3', mRNA sequence.
 DEFINITION

ACCESSION AV661154
 VERSION AV661154.1 GI:9882168
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 929)
 X. X., Huang, J., Xu, Z., Qian, B., Zhu, Z., Yan, Q., Cai, T., Zhang, X., Xiao, H., Qu, J., Liu, F., Huang, Q., Cheng, Z., Li, N., Du, J., Hu, W., Shen, K., Lu, G., Fu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X., Hu, G., Gu, J., Chen, Z. and Han, Z.
 TITLE Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
 MEDLINE 21625106
 COMMENT Contract: Zeguang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
 Tel: 86-21-50801919 (ex. 45)
 Fax: 86-21-50801922
 Email: hanzg@chgc.sh.cn
 This clone is available at CHGC in Shanghai.
 FEATURES
 source 1..660
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="GICPC05"
 /clone_id="GIC"
 /cissue_type="corresponding non cancerous liver tissue"
 /dev_stage="Adult"
 /lab_host="SOLR"
 /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
 BASE COUNT 171 a 167 c 177 g 145 t
 ORIGIN
 Query Match 44.3%; Score 618; DB 10; Length 660;
 Best Local Similarity 98.2%; Pred. No. 1.1e-164;
 Matches 646; Conservative 0; Mismatches 10; Indels 2; Gaps 2;
 479 GCCGACTGATGAAAAGCCAAAGATCTCCAAAGTGTATCAGCCAAATGCCCTTTTGG 538
 1 GCCGACTGATGAAAAGCCAAAGATCTCCAAAGTGTATCAGCCAAATGCCCTTTTGG 60
 539 GAGCAAAATCCCTTACTTCAATGAGACTTACAGAGATCAGTGTGTGTGTGTGTGTGT 598
 61 GAGCAAAATCCCTTACTTCAATGAGACTTACAGAGATCAGTGTGTGTGTGTGTGTGT 120
 599 CCAAGCTCAGCCCTTGAGCTTCAAGAAATGACAGCAATCAGAGCGGCATCAACA 658
 121 CCAAGCTCAGCCCTTGAGCTTCAAGAAATGACAGCAATCAGAGCGGCATCAACA 180
 659 AATGGGTGTCCAAATGAG 718
 181 AATGGGTGTCCAAATGAG 240
 719 ATGAGCTACTGTTCTGT 778
 241 ATGAGCTACTGTTCTGT 300
 779 AGTTCAGCCCTTGAGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 838
 301 AGTTCAGCCCTTGAGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
 839 CAGCATCTATGATGTACAG 898
 361 CAGCATCTATGATGTACAG 420
 899 AGGTGTTGAGTGGCTTCAAG 958

Db 421 AGTGTCTGAGTTCCTTCCTCAAGGTGATGACATCACCATGCTCTCATCTTGCCCAAGC 480
QY 959 CTGAGAGAGCCTGSCCAAGGTGAGAGAGAACTGACCCCAAGGTGTCAGAGAGTGC 1018
Db 481 CTGAGAGAGCCTGSCCAAGGTGAGAGAGAACTGACCCCAAGGTGTCAGAGAGTGC 540
QY 1019 TGGATGAATTGGAGAGAGATGATGCTGTGTCTCACATGCCCCGCTTCCGATTGA-GGAG 1077
Db 541 TGGATGAATTGGAGAGAGATGATGCTGTGTCTCACATGCCCCGCTTCCGATTGAAGGAGC 600
QY 1078 GGCCTTCAGTTGAGAGAGAGCAG-CTGCAAGACATGAGGCTTGTGATCTGTTCAGGCCCT 1134
Db 601 GCTTCAGTTGAGAGAGAGCAGCTGTCTAGACATGAGGCTTGTGATCTGTTCAGGCCCT 658

Search completed: June 27, 2003, 00:05:50
Job time : 1904 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 26, 2003, 16:14:37 ; Search time 3598 Seconds

(without alignments)
11283.606 Million cell updates/sec

Title: US-09-828-592-6

Perfect score: 1395
Sequence: 1 atgcatccatgcatgcatgag.....ccaaccttggttaagtaa 1395Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```
GenEmbl:
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_ay:*
39: em_hlgo_hum:*
40: em_hlgo_mus:*
41: em_hlgo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1395	100.0	1489	6 A22203	A22203 Antithrombi
2	1391.8	99.8	1599	6 A06100	A06100 Synthetic h
3	1391.8	99.8	1599	6 I03102	I03102 Sequence 1
4	1391.8	99.8	1600	6 E00116	E00116 DNA coding
5	1391.8	99.8	1794	17 AF130100	AF130100 Homo sapi
6	1390.2	99.7	1599	6 I02335	I02335 Sequence 1
7	1388.6	99.5	1395	6 I12223	I12223 Sequence 1
8	1385.6	99.3	1392	6 E06109	E06109 DNA encodin
9	1375.6	98.6	1467	9 HUMAT11V	D98832 Human mRNA
10	1114.8	79.9	1751	4 OAA111	X68287 O.aries mRN
11	1063.6	76.2	1509	10 S47225	S47225 antithrombi
12	1063.6	76.2	1672	10 BC019447	BC019447 Mus muscu
13	971.6	69.6	1527	10 AB027238	AB027238 Cavia por
14	702.2	50.3	1362	5 AF411692	AF411692 Chelidra
15	690	49.5	1542	5 AF335570	AF335570 Sphenodon
16	675	48.4	1755	5 AF411691	AF411691 Struthio
17	633.8	45.4	1272	5 S79838	S79838 antithrombi
18	569.4	40.4	1371	5 AF411693	AF411693 Xenopus l
19	564.2	40.4	988	5 BC022309	BC022309 Homo sapi
20	562.2	40.3	1347	5 ASA252153	AU252153 Salmo bal
21	412	29.5	572	4 AF281653	AF281653 Sus scrof
22	401.2	28.8	14206	9 HSAT3	X68793 H.sapiens g
23	401.2	28.8	62125	9 HUMATH3U5	AL691487 Human DNA
24	400.6	28.7	582	9 HUMATH3U5	M21640 Human (dysf
25	399	28.6	452	9 HUMATH3X4	M21640 Human antit
26	398	28.5	16330	9 AP386078	AP386078 Homo sapi
27	372	26.7	432	9 HUMATH3X2	L00186 Human antit
28	372	26.7	604	9 HUMATH3U2	M21637 Human (dysf
29	359.2	25.7	385	9 HUMATH3X3	L00187 Human antit
30	326.4	23.4	20993	10 AB043785	AB043785 Mus muscu
31	276.4	19.8	311	9 HSAT3G1	X03203 Human gene
32	254	18.2	355	6 AX408933	AX408933 Sequence
33	220.8	15.8	398	9 HUMATH3A1	M21643 Human antit
34	220.8	15.8	644	9 HUMATH3U3	M21638 Human (dysf
35	217.8	15.6	1256	5 AF006495	AF006495 Cyprinus
36	191.8	13.7	2490	5 AB026832	AB026832 Fugu rubr
37	182.4	13.1	90216	2 AC113837	AC113837 Rattus no
38	179.8	12.9	238	9 HUMATH3A3	M21645 Human antit
39	179	12.8	230	9 HSAT3G2	X03204 Human antit
40	178.2	12.8	267	9 HUMATH3X6	L00190 Human antit
41	178.2	12.8	324	9 HUMATH3U7	M21642 Human (dysf
42	147.2	10.6	594	11 G16351	G16351 AT3-334P Do
43	144.4	10.4	469	9 HUMATH3A2	M21644 Human antit
44	144.4	10.4	587	9 HUMATH3U4	M21633 Human (dysf
45	144.4	10.4	80780	2 AC020777	AC020777 Homo sapi

ALIGNMENTS

```
RESULT 1
LOCUS A22203 1489 bp mRNA linear PAT 21-JUL-1994
DEFINITION Antithrombin III.
ACCESSION A22203
VERSION A22203.1 GI:583740
KEYWORDS
ORGANISM synthetic construct.
SOURCE synthetic construct.
FEATURES
    source
        location/Qualifiers
            1..1489
                /organism="synthetic construct"
                /db_xref="taxon:326310"
                11..1405
                    gene
```

CDS
/gene="Antithrombin III"
11..1405
/gene="Antithrombin III"
/codon_start=1
/transl_table=11
/protein_id="CAA01592.1"
/db_xref="GI:583741"

BASE COUNT 395 a 372 c 357 g 365 t
ORIGIN
Query Match 100.0%; Score 1395; DB 6; Length 1489;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGATTCATATGTATAGAACTGTAACCTCTGAAAAAGAGATTATCTTTGTC 60
11 ATGATTCATATGTATAGAACTGTAACCTCTGAAAAAGAGATTATCTTTGTC 70
61 TTGCTGCTCATTTGGCTTCTGGGACTGCTGACCTGTCAAGGAGCCCTGTGACATCTG 120
71 TTGCTGCTCATTTGGCTTCTGGGACTGCTGACCTGTCAAGGAGCCCTGTGACATCTG 130
121 ACAGCCAAAGCCGCGGACATTTCCCATGAATCCCATGTGCAATTTACCGCTCCCGAGAG 180
131 ACAGCCAAAGCCGCGGACATTTCCCATGAATCCCATGTGCAATTTACCGCTCCCGAGAG 190
181 AAGGCACTGAGAGATGAGGCTCAGAAAGAGATCCCGAGGAGCCAGCCGCGGTGTC 240
191 AAGGCACTGAGAGATGAGGCTCAGAAAGAGATCCCGAGGAGCCAGCCGCGGTGTC 250
241 TGGGAAGTCTGCAAGGCAATTCGCGCTTGGCTACCACTTTCTATCAGACCTGGCAGAT 300
251 TGGGAAGTCTGCAAGGCAATTCGCGCTTGGCTACCACTTTCTATCAGACCTGGCAGAT 310
301 TCCAAAGATGACATGATTAACATTTTCTGTCAACCCCTGAGATCTCAAGGCTTTTGTCT 360
311 TCCAAAGATGACATGATTAACATTTTCTGTCAACCCCTGAGATCTCAAGGCTTTTGTCT 370
361 ATGACCAAGCTGGGTGCTGTAATGACACCTCCAGCACTGATGAGGATTTAAGTTT 420
371 ATGACCAAGCTGGGTGCTGTAATGACACCTCCAGCACTGATGAGGATTTAAGTTT 430
421 GACACCATATCTGAGAAAACATCTGATCAGATCACTTCTTTGCAAACTGAACTGC 480
431 GACACCATATCTGAGAAAACATCTGATCAGATCACTTCTTTGCAAACTGAACTGC 490
481 CGACTCTATCGAAAAGCCCAAAATCTCTCAAGTTAGTATCAGCAATGCGCTTTTGTGA 540
491 CGACTCTATCGAAAAGCCCAAAATCTCTCAAGTTAGTATCAGCAATGCGCTTTTGTGA 550
541 GACCAATCCCTTACCTTCAATGAGACTCAAGCAATAGAGATGAGTATATGAGGC 600
551 GACCAATCCCTTACCTTCAATGAGACTCAAGCAATAGAGATGAGTATATGAGGC 610
601 AAGCTCCAGCCCTGGAATTCAGAGAAAATGAGAGCAATCCAGAGGCGCATCAACAA 660
611 AAGCTCCAGCCCTGGAATTCAGAGAAAATGAGAGCAATCCAGAGGCGCATCAACAA 670
661 TGGGTGCTCAATTAAGACCGAAGCGGATCAAGGATGATTCCTCGAAGCCATCAAT 720
671 TGGGTGCTCAATTAAGACCGAAGCGGATCAAGGATGATTCCTCGAAGCCATCAAT 730
721 GAGCTCACTGTTCTGGTGTGTTAACACATTTTCAAGGCGCTGTGAAGCAAG 780

Db 721 GAGCTCACTGTTCTGGTGTGTTAACAACATTTTCAAGGCGCTGTGAAGTCAAAG 790
Qy 781 TTCAGCCCTGAGAACACAGAGAGAGATGTTCTTCAAGGCTGTGAGAGTCTGTCTCA 840
Db 791 TTCAGCCCTGAGAACACAGAGAGAGATGTTCTTCAAGGCTGTGAGAGTCTGTCTCA 850
Qy 841 GCATCTATGATGATACAGAGAGAGAGATGTTCTTCAAGGCTGTGAGAGTCTGTCTCA 900
Db 851 GCATCTATGATGATACAGAGAGAGAGATGTTCTTCAAGGCTGTGAGAGTCTGTCTCA 910
Qy 901 GTGCTTGAGTGGCCCTTCAAGGATGATCAATCAAGGCTGTGAGAGTCTGTCTCAAG 960
911 GTGCTTGAGTGGCCCTTCAAGGATGATCAATCAAGGCTGTGAGAGTCTGTCTCAAG 970
Qy 961 GAGAGAGCCCTGAGAACAGAGAGAGATGTTCTTCAAGGCTGTGAGAGTCTGTCTCAAG 1020
971 GAGAGAGCCCTGAGAACAGAGAGAGATGTTCTTCAAGGCTGTGAGAGTCTGTCTCAAG 1030
Qy 1021 GATGATTTGAGAGAGATGATCTGTGTGATCCATGCTCCGCTTCCGATTTGAGAGCGC 1080
Db 1031 GATGATTTGAGAGAGATGATCTGTGTGATCCATGCTCCGCTTCCGATTTGAGAGCGC 1090
Qy 1081 TTGAGTTTGAAGAGAGAGATGATCTGTGTGATCCATGCTCCGCTTCCGATTTGAGAG 1140
1091 TTGAGTTTGAAGAGAGAGATGATCTGTGTGATCCATGCTCCGCTTCCGATTTGAGAG 1150
Qy 1141 TCCAAAGTCCAGGATTTGTTGAGAGAGAGAGATGATCTGTGTGATCCATGCTCCGAT 1200
1151 TCCAAAGTCCAGGATTTGTTGAGAGAGAGAGATGATCTGTGTGATCCATGCTCCGAT 1210
Qy 1201 CATTAAGGATTTCTTGAAGTAAATGAGAGAGAGATGATCTGTGTGATCCATGCTCCGAT 1260
Db 1211 CATTAAGGATTTCTTGAAGTAAATGAGAGAGAGATGATCTGTGTGATCCATGCTCCGAT 1270
Qy 1261 GTGATGCTGGCGCTGCTAAACCCCAAGAGGATCTTCAAGGCGCAAGGCTTTTC 1320
1271 GTGATGCTGGCGCTGCTAAACCCCAAGAGGATCTTCAAGGCGCAAGGCTTTTC 1330
Qy 1321 CTGTTTATTAAGAGAGTCTCTGTAACATTAATCTTCAAGGCGCAAGGCTTTTC 1380
Db 1331 CTGTTTATTAAGAGAGTCTCTGTAACATTAATCTTCAAGGCGCAAGGCTTTTC 1390
Qy 1381 CTTGTGTGAATGA 1395
Db 1391 CTTGTGTGAATGA 1405

RESULT 2

LOCUS A06100 1599 bp mRNA linear PAT 08-JUL-1993
DEFINITION Synthetic human antithrombin III mRNA.
ACCESSION A06100
VERSION A06100.1 GI:412722
KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

CDS

1. 1599
/organism="synthetic construct"
/db_xref="taxon:32630"
121..1515
/codon_start=1
/transl_table=11
/product="synthetic antithrombin III"


```

QY 61 TTGCTGCTCATTTGGCTTCTGGAGACTGCGTGACCTGTGACAGGGAGCCCTGTGACATCTGC 120
DB 181 TTGCTGCTCATTTGGCTTCTGGAGACTGCGTGACCTGTGACAGGGAGCCCTGTGACATCTGC 240
QY 121 ACAGCCAAAGCCGGGAGCAATTCCTCATGAATCCCATGATGATTAACGGCTCCCGGAGAG 180
DB 241 ACAGCCAAAGCCGGGAGCAATTCCTCATGAATCCCATGATGATTAACGGCTCCCGGAGAG 300
QY 181 AAGGCAATGAGAGATGAGGCTTCAGAAACAGAGATCCGAGAGGCCAACACGGCGCTGTG 240
DB 301 AAGGCAATGAGAGATGAGGCTTCAGAAACAGAGATCCGAGAGGCCAACACGGCGCTGTG 360
QY 241 TGGGAATCTGCCAAGGCCAATTCGGCTTGTCTACCACTTTCTATAGCACTCTGGAGAT 300
DB 361 TGGGAATCTGCCAAGGCCAATTCGGCTTGTCTACCACTTTCTATAGCACTCTGGAGAT 420
QY 301 TCCAAAGATGACAAATGATTAATTTCTGTGACCCCTGAGATCTCCACGGCTTTTGTCT 360
DB 421 TCCAAAGATGACAAATGATTAATTTCTGTGACCCCTGAGATCTCCACGGCTTTTGTCT 480
QY 361 ATGACCAAGCTGGGTCCTGTATATGACACCTTCACAGCACTGATGAGGATTTTAAGTT 420
DB 481 ATGACCAAGCTGGGTCCTGTATATGACACCTTCACAGCACTGATGAGGATTTTAAGTT 540
QY 421 GACACCATATCTGAGAAACATCTGATCAATCACTTCTTTTCCCAAGCTGAACTGC 480
DB 541 GACACCATATCTGAGAAACATCTGATCAATCACTTCTTTTCCCAAGCTGAACTGC 600
QY 481 CGACTCTATCGAAACCAAAATCTCTCAAGTATGATCAAGCCAAATGCGCTTTTGTGA 540
DB 601 CGACTCTATCGAAACCAAAATCTCTCAAGTATGATCAAGCCAAATGCGCTTTTGTGA 660
QY 541 GACAAATCCCTTACCTTCAATGAGACCTACACAGACATGATGATGATATATGAGGC 600
DB 661 GACAAATCCCTTACCTTCAATGAGACCTACACAGACATGATGATGATATATGAGGC 720
QY 601 AAGCTCCAGCCCTGAGACTTCAAGAAATGACAGACATCCAGAGCCCTCATCAACAA 660
DB 721 AAGCTCCAGCCCTGAGACTTCAAGAAATGACAGACATCCAGAGCCCTCATCAACAA 780
QY 721 TGGGTCCTCAATTAAGACCGAGCCGATCAACGATGTCATTCCTCGGAGGCCATCAAT 720
DB 841 TGGGTCCTCAATTAAGACCGAGCCGATCAACGATGTCATTCCTCGGAGGCCATCAAT 840
QY 841 GAGCTCACTGTTCTGCTGCTGTTAAACACATTTACTTCAAGGCGCTGTGAAATCAAG 780
DB 901 GAGCTCACTGTTCTGCTGCTGTTAAACACATTTACTTCAAGGCGCTGTGAAATCAAG 900
QY 781 TTACAGCCCTGAGAAACACAGAGAGAACTGTTTCAAGGCTGATGAGAGTCTGTTCA 840
DB 901 TTACAGCCCTGAGAAACACAGAGAGAACTGTTTCAAGGCTGATGAGAGTCTGTTCA 960
QY 841 GCATCATATGATGATGACAGAGAGCAAGTCCGTTATCCGCGCTGCTGAAGCACCAG 900
DB 961 GCATCATATGATGATGACAGAGAGCAAGTCCGTTATCCGCGCTGCTGAAGCACCAG 1020
QY 901 GTGCTTGAAGTTCCTTCAAGGTGATGATCATCAATGATGATGATGATGATGATGAT 960
DB 1021 GTGCTTGAAGTTCCTTCAAGGTGATGATCATCAATGATGATGATGATGATGATGAT 1080
QY 961 GAGAAAGCTGGCCCAAGGTGAGAGAGAACTCAACCCAGAGGTGCTGAGAGAGTGGCTG 1020
DB 1081 GAGAAAGCTGGCCCAAGGTGAGAGAGAACTCAACCCAGAGGTGCTGAGAGAGTGGCTG 1140
QY 1021 GATGATTTGAGAGAGATGATGCTGTGTCTGATCATCCCGCTTCGCAATGAGAGCGC 1080
DB 1141 GATGATTTGAGAGAGATGATGCTGTGTCTGATCATCCCGCTTCGCAATGAGAGCGC 1200
QY 1081 TTACGTTTGAAGAGAGAGCTGCAAGACATGAGGCTTGTGATCTGTTCAAGCCCTGAAAG 1140
DB 1201 TTACGTTTGAAGAGAGAGCTGCAAGACATGAGGCTTGTGATCTGTTCAAGCCCTGAAAG 1260

```

```

QY 1141 TCCAAACTCCAGGATTTGTGAGAAAGCCGAGATGACCTATGCTCAGATGATTC 1200
DB 1261 TCCAAACTCCAGGATTTGTGAGAAAGCCGAGATGACCTATGCTCAGATGATTC 1320
QY 1201 CATAGGATTTCTTGAAGATTAATGAGAGAGCAGTGAAGAGCTGCAAGTACCGCTGT 1260
DB 1321 CATAGGATTTCTTGAAGATTAATGAGAGAGCAGTGAAGAGCTGCAAGTACCGCTGT 1380
QY 1261 GTGATTCCTGGCCGCTGCTTAACCCCAACAGGCTGATCTTCAAGGCCAACAGCCCTTC 1320
DB 1381 GTGATTCCTGGCCGCTGCTTAACCCCAACAGGCTGATCTTCAAGGCCAACAGCCCTTC 1440
QY 1321 CTGCTTTTATPAAGAGAGTCTCTGTAACCTATTTATCTCATGAGCAGAGTACCAAC 1380
DB 1441 CTGCTTTTATPAAGAGAGTCTCTGTAACCTATTTATCTCATGAGCAGAGTACCAAC 1500
QY 1381 CTTGTGTTAAGTAA 1395
DB 1501 CTTGTGTTAAGTAA 1515

RESULT 4
E00116
LOCUS E00116 1600 bp RNA linear PAT 29-SEP-1997
DEFINITION DNA coding of human anti-chrombin III.
ACCESSION E00116 GI:2168417
VERSION E00116.1
KEYWORDS JP 1983162529-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1600)
AUTHORS Suizan,K.B. and Richiyoado,E.R.
TITLE HUMAN ANTITHROMBOTIC III
JOURNAL Patient: JP 1983162529-A 1 27-SEP-1983;
GENENTECH INC
COMMENT OS human
PN JP 1983162529-A/1
PD 27-SEP-1983
PF 03-MAR-1983 JP 1983033900
PR 03-MAR-1982 GB 82 8206262, 30-JUL-1982 US 82 403600 PI
SUZAN KURAAKU BOTSUKU, RICHIIYADO EMU ROON
PC A6IK37/04,A6IK35/74,A6IK37/64,C07H21/00,C12N15/00; CC
strandness: Double;
CC topology: linear;
CC hypothetical: No;
CC anti-sense: No;
*source: tissue=liver;
FH key Location/Qualifiers
FT CDS 121..1515
FT mat_peptide 217..1512 /product='human anti-chrombin III' FT
FT sig_peptide 121..216 /product='mature human anti-chrombin III' FT
FEATURES
source Location/Qualifiers
BASE COUNT 422 a 409 c 375 g 394 t
ORIGIN
Query Match 99.8%; Score 1391.8; DB 6; Length 1600;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1393; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1. ATGATTTCCATNGATGAGAACTGTAACCTCTGAAAGAGAGGTTATCTTTGTC 60
DB 121 ATGATTTCCATNGATGAGAACTGTAACCTCTGAAAGAGAGGTTATCTTTGTC 180
QY 61 TTGCTGCTCATTTGGCTTCTGGAGACTGCGTGACCTGTGACAGGGAGCCCTGTGACATCTGC 120

```

```

Db 181 TTGCTGCTCATTTGGCTTCTGGGACTGCGTGAACCTGTCAACGGGAGCCCTGTGGACATCTGC 240
Oy 121 AAGCCAAAGCGCGGGACATTTCCCATGATCCCATGATGATTTACCGCTCCCGGAGAG 180
Db 241 AAGCCAAAGCGCGGGACATTTCCCATGATGATGATTTACCGCTCCCGGAGAG 300
Oy 181 AAGCCAAAGCGCGGGACATTTCCCATGATGATGATTTACCGCTCCCGGAGAG 240
Db 301 AAGCCAAAGCGCGGGACATTTCCCATGATGATGATTTACCGCTCCCGGAGAG 360
Oy 241 TGGGAATCTGTCAGAGGCCAATTTCCCGCTTGTCTACCATTTCTATCAGACCTGGCAGAT 300
Db 361 TGGGAATCTGTCAGAGGCCAATTTCCCGCTTGTCTACCATTTCTATCAGACCTGGCAGAT 420
Oy 301 TCCGAATGACATGATTAACATTTCCGTCACCCCTGAGATGATCCAGCGCTTGTGCT 360
Db 421 TCCGAATGACATGATTAACATTTCCGTCACCCCTGAGATGATCCAGCGCTTGTGCT 480
Oy 361 ATGACCAAGCTGGGTGCTGTAATGACACCCCTCCAGACATGATGAGGATTTAAAGTTT 420
Db 481 ATGACCAAGCTGGGTGCTGTAATGACACCCCTCCAGACATGATGAGGATTTAAAGTTT 540
Oy 421 GACACCATATCTGAGAAAAATCTGATGATGATCCATCTTCTTTGCCAACTGAACTGC 480
Db 541 GACACCATATCTGAGAAAAATCTGATGATGATCCATCTTCTTTGCCAACTGAACTGC 600
Oy 481 CGACTCTATCGAAAAAGCCAAATCCTCCAAAGTATGATGAGGCAATCGCCTTTTGA 540
Db 601 CGACTCTATCGAAAAAGCCAAATCCTCCAAAGTATGATGAGGCAATCGCCTTTTGA 660
Oy 541 GACAAATCCTTACCTTCAATGAGACTACAGAGACATCAGTGAATGATGATGAGACC 600
Db 661 GACAAATCCTTACCTTCAATGAGACTACAGAGACATCAGTGAATGATGATGAGACC 720
Oy 601 AAGCTCCAGCCCTTGAATTTCAAGAAAAATGACAGCAATCCAGCGCCATCAACAA 660
Db 721 AAGCTCCAGCCCTTGAATTTCAAGAAAAATGACAGCAATCCAGCGCCATCAACAA 780
Oy 661 TGGGTCATTAAGACCGAAGCGGATCAACCGATGATTCCTCCGAAAGCCATCAT 720
Db 781 TGGGTCATTAAGACCGAAGCGGATCAACCGATGATTCCTCCGAAAGCCATCAT 840
Oy 721 GAGCTCACTGTTCTGCTGCTGTTAAACATTTTCAAGAGGCTGTGAAAGTCAAG 780
Db 841 GAGCTCACTGTTCTGCTGCTGTTAAACATTTTCAAGAGGCTGTGAAAGTCAAG 900
Oy 781 TTCAGCCCTGAGAACACAGAGGAACTGTTCTCAAGAGCTGATGAGAGTGTGTTCA 840
Db 901 TTCAGCCCTGAGAACACAGAGGAACTGTTCTCAAGAGCTGATGAGAGTGTGTTCA 960
Oy 841 GCATCTATGATGATCAAGAGGAAAGTTCCTGTTACCGCGGTGGCTGAAGGCAACAG 900
Db 961 GCATCTATGATGATCAAGAGGAAAGTTCCTGTTACCGCGGTGGCTGAAGGCAACAG 1020
Oy 901 GTGCTTGAAGTGGCCCTTCAAGAGTATGATCAACATGATGATCTGATCTTGGCCAGGCT 960
Db 1021 GTGCTTGAAGTGGCCCTTCAAGAGTATGATCAACATGATGATCTGATCTTGGCCAGGCT 1080
Oy 961 GAGAAAGCTTGGCCCAAGGTGAGAAAGAACTCACCCCAAGAGTGTGAGAGTGGCTG 1020
Db 1081 GAGAAAGCTTGGCCCAAGGTGAGAAAGAACTCACCCCAAGAGTGTGAGAGTGGCTG 1140
Oy 1021 GATGAATTTGAGAGGATGATGCTGTGTGTCACATGCCCCCGTTCCGATTTGAGAGCGC 1080
Db 1141 GATGAATTTGAGAGGATGATGCTGTGTGTCACATGCCCCCGTTCCGATTTGAGAGCGC 1200
Oy 1081 TTCAGTTTGAAGAGCAGCTGCAAGACATGAGGCTTGTGATCTGTTCAAGCCCTGAAG 1140
Db 1201 TTCAGTTTGAAGAGCAGCTGCAAGACATGAGGCTTGTGATCTGTTCAAGCCCTGAAG 1260
Oy 1141 TCCAAACTCCAGGATATGTTGTCAGAGAGCCGAGATGATCTTATGTTCTCAGATCATT 1200
Db 1261 TCCAAACTCCAGGATATGTTGTCAGAGAGCCGAGATGATCTTATGTTCTCAGATCATT 1320

```

```

Oy 1201 CATAGGCAATTTCTTGAGGTAAATGAGAAAGGAGTGAAGCAGCTGCAAGTACCGCTGTT 1260
Db 1321 CATAGGCAATTTCTTGAGGTAAATGAGAAAGGAGTGAAGCAGCTGCAAGTACCGCTGTT 1380
Oy 1261 GTGATTCCTGGCCCTTGTGCTAAACCCCAAGGCTGATCTTTCAAGGCCCAAGGCTTTTC 1320
Db 1381 GTGATTCCTGGCCCTTGTGCTAAACCCCAAGGCTGATCTTTCAAGGCCCAAGGCTTTTC 1440
Oy 1321 CTGCTTTTATTAAGAAAGTTCCTTGAACACTATTTATCTTCATGAGGCGAGATGCCAAC 1380
Db 1441 CTGCTTTTATTAAGAAAGTTCCTTGAACACTATTTATCTTCATGAGGCGAGATGCCAAC 1500
Oy 1381 CCTGTGTTAGTAA 1395
Db 1501 CCTGTGTTAGTAA 1515

RESULT 5
AF130100 standard; RNA; HUM; 1794 BP.
ID AF130100,
AC AF130100,
XX AF130100.1
SV AF130100.1
XX
DT 03-DEC-2000 (Rel. 66, Created)
DT 09-MAY-2001 (Rel. 67, Last updated, Version 2)
DE Homo sapiens clone FLB1413 PRO0309 mRNA, complete cds.
XX
XX Homo sapiens (human)
XX OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
XX OC Eutheria; Primates; Catarrhini; Homiidae; Homo.
XX
XX [1]
XX RN 1-1794
XX RP Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y., Xu W.,
XX RA Gao F., Liu M., He F.;
XX RT "Functional prediction of the coding sequences of 75 new genes deduced by
XX RL analysis of cDNA clones from human fetal liver";
XX
XX [2]
XX RN 1-1794
XX RP Zhang C., Yu Y., Zhang S., Wei H., Bi J., Zhou G., Dong C., Zai Y., Xu W.,
XX RA Gao F., Liu M., He F.;
XX RT Submitted (23-FEB-1999) to the EMBL/Genbank/DBJ databases.
XX RL Department of Experimental Hematology, Institute of Radiation Medicine,
XX RL Beijing Taiding Road 27, Beijing, Beijing 100850, P. R. China
XX
XX SWISS-PROT; P01008; ANT3_HUMAN.
XX
XX Key location/Qualifiers
XX FH 1. 1794
XX FT /db_xref="taxon:9606"
XX FT /organism="Homo sapiens"
XX FT /clone="FLB1413"
XX FT /rissue_type="liver"
XX FT /dev_stage="fetus"
XX FT repeat_region
XX FT 16. 137
XX FT /rpe_family="Alu"
XX FT /rpe_type=DISPERSED
XX FT 234. 1628
XX FT /codon_start=1
XX FT /db_xref="SWISS-PROT:P01008"
XX FT /evidence=NOT_EXPERIMENTAL
XX FT /note="predicted protein of H00309"
XX FT /product="PRO0309"

```

```

FT      /protein_id="AAG35525.1"
FT      /translation="MYSNVIQVTSGRKRVYLISLLIFPMDCVTCGSPVDICTAKPR
FT      DIIPNMCYRSPKATDEBSEOKIPATNRNVELSKANSRPFATTEYQHLASKND
FT      NNIIFSPISITAPFAMTKLACNDILOMEVFKPDIISEKTSDDIHFFKLANCLY
FT      RANKSKSLVANSRFLGDSLPENETYOISLGVAKIOPIDFKENASOSPAINKW
FT      SKTEGRITDVI PSEAINELVIVNVTIYFKGIMSKRSPENRKEIYPKADGSCSA
FT      SMYVQGRKRRYRVAGTQVLELPFKGDITWVLIIPREKSLAKREKLEPVLQEM
FT      DELSEMLVIVHFRIRIDGFSLEQLOMDGLVDFSPKSKLPGIIVAGRDULVSDA
FT      FPKAFLEVNEBGESEAASTAIVIAGRSLNPNRVTFKRNPFVIFIREVLNLTIFMGRV
FT      ANPCVR"
XX
XX
SO      Sequence 1794 bp; 540 A; 429 C; 420 G; 405 T; 0 other;

Query Match      99.8%; Score 1391.8; DB 17; Length 1794;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1393; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  ATGTATTCGAATGTGATAGGAATCTGAACCTCTGGAAGAAAGAGTTATCTTTGTCC 60
DB      234  ATGTATTCGAATGTGATAGGAATCTGAACCTCTGGAAGAAAGAGTTATCTTTGTCC 293
QY      61  TTGCTGCTCATTTGGCTTTGGGACTGCGTGAACCTGTCAAGGAGCCCTGTGACATCTGC 120
DB      294  TTGCTGCTCATTTGGCTTTGGGACTGCGTGAACCTGTCAAGGAGCCCTGTGACATCTGC 353
QY      121  ACAGCCAAAGCCGGGACATTTCCCATGATATCCCATGTGATATTCGCTCCCGAGAAAG 180
DB      354  ACAGCCAAAGCCGGGACATTTCCCATGATATCCCATGTGATATTCGCTCCCGAGAAAG 413
QY      181  AAGGCAATCTGAGATAGAGGCTCAGAAACGAAATCCCGAGGCGACCAACCGGCGTGC 240
DB      414  AAGGCAATCTGAGATAGAGGCTCAGAAACGAAATCCCGAGGCGACCAACCGGCGTGC 473
QY      241  TGGGAATCTGTCAAGGCAATTTCCGCTTTGCTACCACTTTCTATGAGCACTGTGAGAT 300
DB      474  TGGGAATCTGTCAAGGCAATTTCCGCTTTGCTACCACTTTCTATGAGCACTGTGAGAT 533
QY      301  TCCAAAGATGACATGATTAATTTCTCTGTCACCCCTGATGATCTTCACGCGCTTTGCT 360
DB      534  TCCAAAGATGACATGATTAATTTCTCTGTCACCCCTGATGATCTTCACGCGCTTTGCT 593
QY      361  ATGACCAAGCTGGGTGCTGTATGACACCTCCAGAACATGATGAGGATTTAGATT 420
DB      594  ATGACCAAGCTGGGTGCTGTATGACACCTCCAGAACATGATGAGGATTTAGATT 653
QY      421  GACACCATATCTGAGAAAACATCTGATCAGATCCACTTTCTTTGCCAACTGAACTGC 480
DB      654  GACACCATATCTGAGAAAACATCTGATCAGATCCACTTTCTTTGCCAACTGAACTGC 713
QY      481  CGACTCTATGAGAAAACGCAAAATCTCTCAAGTTAGTATGAGCAATGCGCTTTTGGCA 540
DB      714  CGACTCTATGAGAAAACGCAAAATCTCTCAAGTTAGTATGAGCAATGCGCTTTTGGCA 773
QY      541  GACAAATCCCTTCACTTCAATGAGACTTACCAAGATCAGATGAGTGGATATGAGGC 600
DB      774  GACAAATCCCTTCACTTCAATGAGACTTACCAAGATCAGATGAGTGGATATGAGGC 833
QY      601  AAGCTCCAGCCCTGTGACTTCAAGAAAATGAGAGCAATCCAGAGCGGCATCAACAA 660
DB      834  AAGCTCCAGCCCTGTGACTTCAAGAAAATGAGAGCAATCCAGAGCGGCATCAACAA 893
QY      661  TGGGTGCTCAATTAAGACCGAAGCCGATATCAGATGATCTTCCCTCGAAGCATCAAT 720
DB      894  TGGGTGCTCAATTAAGACCGAAGCCGATATCAGATGATCTTCCCTCGAAGCATCAAT 953
QY      721  GAGCTACTGTTCTGGGTGCTGTTAACACCATTTACTTCAAGGGCCGTGGAAGTCAAG 780
DB      954  GAGCTACTGTTCTGGGTGCTGTTAACACCATTTACTTCAAGGGCCGTGGAAGTCAAG 1013
QY      781  TTCAGCCCTGAGAACCAAGAGAGAACTGTTTCAACAGGCTGATGAGAGTGTGTTCA 840
DB      1014  TTCAGCCCTGAGAACCAAGAGAGAACTGTTTCAACAGGCTGATGAGAGTGTGTTCA 1073

```

```

QY      841  GCATCTATGATGATACAGAGAGCAAGTCCGTTATCGCCCGCTGAGGACCCAG 900
DB      1074  GCATCTATGATGATACAGAGAGCAAGTCCGTTATCGCCCGCTGAGGACCCAG 1133
QY      901  GGGCTTGAAGTCCCTTCAAGGTATGACATACCATGTGTCCTCATCTTGCCCAAGCT 960
DB      1134  GGGCTTGAAGTCCCTTCAAGGTATGACATACCATGTGTCCTCATCTTGCCCAAGCT 1193
QY      961  GAGAGAGCTCGGCCAAGGTGAGAGGAAGCACTCACCCAGAGTGTCTCAGAGTGGCTG 1020
DB      1194  GAGAGAGCTCGGCCAAGGTGAGAGGAAGCACTCACCCAGAGTGTCTCAGAGTGGCTG 1253
QY      1021  GATGATTTGAGAGATATGATGTGTGTCACATACCCCGCTTCCGATTTGAGAGAGCGC 1080
DB      1254  GATGATTTGAGAGATATGATGTGTGTCACATACCCCGCTTCCGATTTGAGAGAGCGC 1313
QY      1081  TTCAGTTGAAGAGACAGCTGCAACATGAGGCGCTTTCGATCTTGTCAAGCCCTGAAAAG 1140
DB      1314  TTCAGTTGAAGAGACAGCTGCAACATGAGGCGCTTTCGATCTTGTCAAGCCCTGAAAAG 1373
QY      1141  TCCAAATCCCGAGTATTTGTTCAGAGAGCCGAGATGACCTCTATGTCTCAGATCATTC 1200
DB      1374  TCCAAATCCCGAGTATTTGTTCAGAGAGCCGAGATGACCTCTATGTCTCAGATCATTC 1433
QY      1201  CATTAAGCATTTCTTGAGTAAATTAAGAAGCAGTGAAGCAGTGCATACCTGCTGT 1260
DB      1434  CATTAAGCATTTCTTGAGTAAATTAAGAAGCAGTGAAGCAGTGCATACCTGCTGT 1493
QY      1261  GTGATTTGCTGGCGCTTTCGCTAAACCCCAAGAGGAGTCACTTTCAAGGCCCAAGCGCTTTC 1320
DB      1494  GTGATTTGCTGGCGCTTTCGCTAAACCCCAAGAGGAGTCACTTTCAAGGCCCAAGCGCTTTC 1553
QY      1321  CTGCTTTTATTAAGAAAGTCTCTGTAACATATATCTTATGAGGAGAGTACCAAC 1380
DB      1554  CTGCTTTTATTAAGAAAGTCTCTGTAACATATATCTTATGAGGAGAGTACCAAC 1613
QY      1381  CCTTGCTTAAGTAA 1395
DB      1614  CCTTGCTTAAGTAA 1628

RESULT 6
LOCUS      102335 1599 bp ss-DNA linear PAT 18-MAY-1993
DEFINITION Sequence 1 from Patent US 4517294.
ACCESSION 102335
VERSION 102335.1 GI:268048
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1599)
AUTHORS Bock,S.C. and Lawn,R.M.
TITLE Human antithrombin III
JOURNAL Patent: US 4517294-A 1 14-MAY-1995;
Genentech, Inc.; South San Francisco, CA
FEATURES
source 1..1599
/organism="unknown"
BASE COUNT 421 a 408 c 375 g 395 t
ORIGIN
Query Match 99.7%; Score 1390.2; DB 6; Length 1599;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1392; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  ATGTATTCGAATGTGATAGGAATCTGAACCTCTGGAAGAAAGAGTTATCTTTGTCC 60
DB      121  ATGTATTCGAATGTGATAGGAATCTGAACCTCTGGAAGAAAGAGTTATCTTTGTCC 180
QY      61  TTGCTGCTCATTTGGCTTTGGGACTGCGTGAACCTGTCAAGGAGCCCTGTGACATCTGC 120
DB      181  TTGCTGCTCATTTGGCTTTGGGACTGCGTGAACCTGTCAAGGAGCCCTGTGACATCTGC 240

```

Qy	121	TCAGCCAGACCGCGGGACATTTCCATGATCCCATGTGATTTACCGCTCCCGGGAG	180
Db	241	ACAGCCAGACCGCGGGACATTTCCATGATTTCCATGTGATTTACCGCTCCCGGGAG	300
Qy	181	AAGGCAACTGAGATGAGGGCTCAGAACAGAGATCCCGAGGCCACCAACCGCGGTTC	240
Db	301	AAGGCAACTGAGATGAGGGCTCAGAACAGAGATCCCGAGGCCACCAACCGCGGTTC	360
Qy	241	TGGGAACCTGTCCAAAGGCCAATTTCCCGCTTGTCTACCACTTTCTATCAGACTGGCAGAT	300
Db	361	TGGGAACCTGTCCAAAGGCCAATTTCCCGCTTGTCTACCACTTTCTATCAGACTGGCAGAT	420
Qy	301	TCCAAGATGACAAATGATTAACATTTTTCCTGTCAACCCTGAGATCTCCAGCGCTTTGGCT	360
Db	421	TCCAAGATGACAAATGATTAACATTTTTCCTGTCAACCCTGAGATCTCCAGCGCTTTGGCT	480
Qy	361	ATGACCAAGCTGGGTGCTCTGTATATGACACCTCCAGCAACTGATGAGAGTATTAAGTTT	420
Db	481	ATGACCAAGCTGGGTGCTCTGTATATGACACCTCCAGCAACTGATGAGAGTATTAAGTTT	540
Qy	421	GACCCCATATCTGAGAAAAACATCTGATCAGATCCACTCTTCTTGGCCAAATCGAACCTGC	480
Db	541	GACCCCATATCTGAGAAAAACATCTGATCAGATCCACTCTTCTTGGCCAAATCGAACCTGC	600
Qy	481	CGACTCTATCGAAAAAGCCAACAAATCTCTCCAGTTATGATTCAGCCAAATGCGCTTTTGTGA	540
Db	601	CGACTCTATCGAAAAAGCCAACAAATCTCTCCAGTTATGATTCAGCCAAATGCGCTTTTGTGA	660
Qy	541	GACCAATCTCCTTAACCTTCAATGAGACTTACCGAGACTATGATGATTTGATATATGAGCC	600
Db	661	GACCAATCTCCTTAACCTTCAATGAGACTTACCGAGACTATGATGATTTGATATATGAGCC	720
Qy	601	AAGCTCCAGCCCTCGACTTCAGAGAAAAATGAGAGCAATCCAGAGCGGCATTCACAA	660
Db	721	AAGCTCCAGCCCTCGACTTCAGAGAAAAATGAGAGCAATCCAGAGCGGCATTCACAA	780
Qy	661	TGGGTGTCCAATAGACCGAAGGCCGAATCACCGATGTCTATTCCTCGAAGCCATCAAT	720
Db	781	TGGGTGTCCAATAGACCGAAGGCCGAATCACCGATGTCTATTCCTCGAAGCCATCAAT	840
Qy	721	GAGCTCACTGTTCTGGTGCTGTTTAAACACATTTTACTTCAAGGGCTGTGGAAGTCAAG	780
Db	841	GAGCTCACTGTTCTGGTGCTGTTTAAACACATTTTACTTCAAGGGCTGTGGAAGTCAAG	900
Qy	781	TTTCAGCCCTTGAGAACACAGAGAGAACTGTTCTACAGGCTGATGAGAGTCTGTCTA	840
Db	901	TTTCAGCCCTTGAGAACACAGAGAGAACTGTTCTACAGGCTGATGAGAGTCTGTCTA	960
Qy	841	GCATCTATGATGTATCCAGAGAGGCAATTCGGTATCGGCGCGTGGCTAAGGCAACCCAG	900
Db	961	GCATCTATGATGTATCCAGAGAGGCAATTCGGTATCGGCGCGTGGCTAAGGCAACCCAG	1020
Qy	901	GTGCTTGAGTTGGCCCTTCAAGAGTATGACATCACCATGGTCTCTATCTTGGCCAAAGCT	960
Db	1021	GTGCTTGAGTTGGCCCTTCAAGAGTATGACATCACCATGGTCTCTATCTTGGCCAAAGCT	1080
Qy	961	GAGAAAGACCTTGGCCAAAGGTGAGAGAACTCACCCAGAGAGTCTCTCAGAGATGGCTG	1020
Db	1081	GAGAAAGACCTTGGCCAAAGGTGAGAGAACTCACCCAGAGAGTCTCTCAGAGATGGCTG	1140
Qy	1021	GATGAAATGAGAGAAATGATGTGTGTGTCTCAATGCCCCGCTTCCGCAATTAGAGACGGC	1080
Db	1141	GATGAAATGAGAGAAATGATGTGTGTGTCTCAATGCCCCGCTTCCGCAATTAGAGACGGC	1200
Qy	1081	TTTCAGTTTGAAGAGACAGCTGCACAAACATGGGCTTGTGATCTGTTCAGCCTTGAAAG	1140
Db	1201	TTTCAGTTTGAAGAGACAGCTGCACAAACATGGGCTTGTGATCTGTTCAGCCTTGAAAG	1260
Qy	1141	TTCCAAACTTCCAGAGTATGTGTGCAAGAGGCGGAGATGACTCTATATGTCTCAATGATTC	1200
Db	1261	TTCCAAACTTCCAGAGTATGTGTGCAAGAGGCGGAGATGACTCTATATGTCTCAATGATTC	1320

Qy	1201	CATAGAGCATTTCTTGAGGTAAATAGAGAGGAGGTGAAGAGCTGCAGATACCGTGTT	1260
Db	1321	CATAGGCAATTTCTTGAGGTAAATAGAGAGGAGGTGAAGAGCTGCAGATACCGTGTT	1380
Qy	1261	GTGATGTGGCCGTTGCGCTAAACCCCAACAGGGTGACTTTCAAGGCCAACAGGCTTTTC	1320
Db	1381	GTGATGTGGCCGTTGCGCTAAACCCCAACAGGGTGACTTTCAAGGCCAACAGGCTTTTC	1440
Qy	1321	CTGGTTTTTAAAGAGAAAGTTCCCTCTGACACTATTTATCTTCATGGGCAAGATGCCAAC	1380
Db	1441	CTGGTTTTTAAAGAGAAAGTTCCCTCTGACACTATTTATCTTCATGGGCAAGATGCCAAC	1500
Qy	1381	CCTGTGTAAAGTAA 1395	
Db	1501	CCTGTGTAAAGTAA 1515	
RESULT 7.			
LOCUS	112223	1395 bp	DNA linear PAT 26-JUL-1995
DEFINITION	Sequence 1 from patent US 5420252.		
ACCESSION	112223		
VERSION	112223.1	GI:909721.	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 1395)		
	Kato,H., Yoshitake,S., Suzuki,S., Suzuki,N., Seto,T., Nagaoka,N.		
	and Mizui,Y.		
TITLE	Human antithrombin III mutants		
JOURNAL	Patent: US 5420252-A 1 30-MAY-1995;		
FEATURES	Location/Qualifiers		
source	1..1395		
	/organism="unknown"		
BASE COUNT	366 a 353 c 343 g 331 t		
ORIGIN			
Query Match	99.5%; Score 1388.6; DB 6; Length 1395;		
Best Local Similarity	99.7%; Pred. No. 0;		
Matches 1391;	Conservative 0; Mismatches 4; Indels 0; Gaps 0;		
Qy	1	ATGATATTCATGTGATGAGAACTGTAACTCTCTGGAAAAGAGGTTATCTTTGTCC	60
Db	1	ATGATATTCATGTGATGAGAACTGTAACTCTCTGGAAAAGAGGTTATCTTTGTCC	60
Qy	61	TTGCGTCATATGGCTTCTGGGACATGGGTGACCTGTACACGGGAGCCCTGTGGACATCTGC	120
Db	61	TTGCGTCATATGGCTTCTGGGACATGGGTGACCTGTACACGGGAGCCCTGTGGACATCTGC	120
Qy	121	ACAGCCAAAGCCGCGGACATTCGCCATGATCCCATGTGCACTTTCACGCTCCCGAGAG	180
Db	121	ACAGCCAAAGCCGCGGACATTCGCCATGATCCCATGTGCACTTTCACGCTCCCGAGAG	180
Qy	181	AAGCAACTGAGATGAGAGGCTCAGAAACAGAAATCCCGAGGCCAACCAACCGCGTGTCTC	240
Db	181	AAGCAACTGAGATGAGAGGCTCAGAAACAGAAATCCCGAGGCCAACCAACCGCGTGTCTC	240
Qy	241	TGGGAACGTGCCAAGGCCAATTCGCCGTTTGCCTACACTTCTATACAGCACTGTGGCAGAT	300
Db	241	TGGGAACGTGCCAAGGCCAATTCGCCGTTTGCCTACACTTCTATACAGCACTGTGGCAGAT	300
Qy	301	TCCAAAGATGACATGATTAACATTTTCTCTGCACCCCTGAGTATCTCCAGCGCTTTTGCT	360
Db	301	TCCAAAGATGACATGATTAACATTTTCTCTGCACCCCTGAGTATCTCCAGCGCTTTTGCT	360
Qy	361	ATGACCAAGCTGGGTGCTGTAAATGACACCTCCAGCAACTGATGAGAGTATTTAAATTT	420
Db	361	ATGACCAAGCTGGGTGCTGTAAATGACACCTCCAGCAACTGATGAGAGTATTTAAATTT	420
Qy	421	GACACCAATCTTGAGAAACATCTGATCAATTCACATTTCTCTTGGCAACATGAATCTGC	480
Db	421	GACACCAATCTTGAGAAACATCTGATCAATTCACATTTCTCTTGGCAACATGAATCTGC	480

QY 481 CGACTCTATCGAAAAGCAAAATCTCCAGTAGTATCAGCCAAATCGCTTTTGG 540
 DB 481 CGACTCTATCGAAAAGCAAAATCTCCAGTAGTATCAGCCAAATCGCTTTTGG 540
 QY 541 GACAAATCCCTTACCTTCAATGAGACTTACAGAGAACTCAATGATTTGGTATGAGCC 600
 DB 541 GACAAATCCCTTACCTTCAATGAGACTTACAGAGAACTCAATGATTTGGTATGAGCC 600
 QY 601 AAGCTCCAGCCCTGAGACTTCAAGAAAATGACAGCAATCCAGAGCCCATCAACAA 660
 DB 601 AAGCTCCAGCCCTGAGACTTCAAGAAAATGACAGCAATCCAGAGCCCATCAACAA 660
 QY 661 TGGGTGTCATTAAGACCCGAGGCGCAATCCAGATGCTATTCCTCGAAGCATCAAT 720
 DB 661 TGGGTGTCATTAAGACCCGAGGCGCAATCCAGATGCTATTCCTCGAAGCATCAAT 720
 QY 721 GAGCTCAGCTTCTGTGTGTGTTACACCATTTACTTCAAGGCGCTGTGGAAGTCAAG 780
 DB 721 GAGCTCAGCTTCTGTGTGTGTTACACCATTTACTTCAAGGCGCTGTGGAAGTCAAG 780
 QY 781 TTCAGCCCTGAGAACACAAAGAAAGAACTGTTCTACAGGCTGATGAGAGTCTGTCA 840
 DB 781 TTCAGCCCTGAGAACACAAAGAAAGAACTGTTCTACAGGCTGATGAGAGTCTGTCA 840
 QY 841 GCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 DB 841 GCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
 QY 901 GGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 DB 901 GGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 QY 961 GAGAAAGCCTGCGCAAGGTGAGAAAGAACTCAGCCAGAGGTGCTGAGAGAGTGGCTG 1020
 DB 961 GAGAAAGCCTGCGCAAGGTGAGAAAGAACTCAGCCAGAGGTGCTGAGAGAGTGGCTG 1020
 QY 1021 GATGAATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
 DB 1021 GATGAATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
 QY 1081 TTCAGTTTAAAG 1140
 DB 1081 TTCAGTTTAAAG 1140
 QY 1141 TCCAAATCCCAAGATTTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
 DB 1141 TCCAAATCCCAAGATTTGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
 QY 1201 CATAGGCAATTTCTTGAAGTAAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
 DB 1201 CATAGGCAATTTCTTGAAGTAAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
 QY 1261 GATGATGAG 1320
 DB 1261 GATGATGAG 1320
 QY 1321 CTGGTTTAAAG 1380
 DB 1321 CTGGTTTAAAG 1380
 QY 1381 CCTTGTGTTAAGTAA 1395
 DB 1381 CCTTGTGTTAAGTAA 1395

RESULT 8
 E06109 1392 bp DNA linear PAT 29-SBP-1997
 LOCUS E06109
 DEFINITION DNA encoding antithrombin III derivative.
 ACCESSION E06109
 VERSION E06109.1 GI:2174296
 KEYWORDS JP 1993339292-A/1.

SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 1392)
 AUTHORS Kato, H., Yoshitake, S., Suzuki, T., Suzuki, N., Seto, T., Nagaoka, N.
 TITLE HUMAN ANTITHROMBIN III MUTANT
 JOURNAL Patent: JP 1993339292-A 1 21-DEC-1993;
 EISAI CO LTD
 COMMENT OS Artificial gene
 OC Artificial sequence; Genes.
 PN JP 1993339292-A/1
 PD 21-DEC-1993
 PF 22-FEB-1993 JP 1993031855
 PR 10-APR-1992 JP 92P 90488
 PI KATO HIROYUKI, YOSHITAKE SHINJI, SUZUKI TAKESHI, SUZUKI NOBORU, PI SETO TOSHIO, NAGAOKA NAOKA, MIZUI YOSHIHARU
 PC C07K13/00, C12N5/10, C12N15/15, C12P21/02, (C12P21/02, C12R1:91);
 CC strandedness: Double;
 CC topology: linear;
 CC Feature is identified by experimental;
 CC Feature is identified by similarity;
 FH Key Location/Qualifiers
 FT CDS 1..1395
 FT sig_peptide 1..96
 FT mat_peptide 97..1392.
 FEATURES Location/Qualifiers
 source 1..1392
 /organism="synthetic construct"
 /db_xref="taxon:32630" 330 t
 BASE COUNT 366 a 353 c 343 g 330 t
 ORIGIN
 Query Match 99.3%; Score 1385.6; DB 6; Length 1392;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1388; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 ATGATTCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 60
 DB 1 ATGATTCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 60
 QY 61 TTGCTGCTCATTTGCTTCTGAGAGTGGTGAATCTGTCAAGGAGCCCTGTGACATCTGC 120
 DB 61 TTGCTGCTCATTTGCTTCTGAGAGTGGTGAATCTGTGTCAAGGAGCCCTGTGACATCTGC 120
 QY 121 ACAGCCAGCGCGGAGACATTCCTCCATGATCCATGATCCATTTACCGCTCCCGGAGAAG 180
 DB 121 ACAGCCAGCGCGGAGACATTCCTCCATGATCCATGATCCATTTACCGCTCCCGGAGAAG 180
 QY 181 AAGGCACTGAGAGATGAGAGGCTCAGACAGAAAGATCCGAGAGCCAGCCAGCGGTGTC 240
 DB 181 AAGGCACTGAGAGATGAGAGGCTCAGACAGAAAGATCCGAGAGCCAGCCAGCGGTGTC 240
 QY 241 TGGGAATCTGTCAGAGCCCAATTCCTGCTTGTGCTACACTTTCTATCAGACCTGGCAAT 300
 DB 241 TGGGAATCTGTCAGAGCCCAATTCCTGCTTGTGCTACACTTTCTATCAGACCTGGCAAT 300
 QY 301 TCCAGAAATGACAAATGATTAATTTCCGTCACCCCTGAGATCTCCAGGCTTTTGTCT 360
 DB 301 TCCAGAAATGACAAATGATTAATTTCCGTCACCCCTGAGATCTCCAGGCTTTTGTCT 360
 QY 361 ATGACCAAGCTGGGTGCTGTGATATGACACCTCCAGCACTGATGAGAGATTTAAGTTT 420
 DB 361 ATGACCAAGCTGGGTGCTGTGATATGACACCTCCAGCACTGATGAGAGATTTAAGTTT 420
 QY 421 GACACCAATCTGAGAAAACATCTGATATGATCCACTTCTTTGGCCAACTGAATGTC 480
 DB 421 GACACCAATCTGAGAAAACATCTGATATGATCCACTTCTTTGGCCAACTGAATGTC 480
 QY 481 CGACTCTATCGAAAAGCAAAATCTCCAGTAGTATCAGCCAAATCGCTTTTGG 540

Db 481 CGACTCTATCGAAGAAACCAACAAATCTCTCAAGTTAGTATCAGCCAAATCGCTTTTGG 540
 Qy 541 GACAAATCCCTTACCTTGAATGAGACCTACGAGCATCAGATGTTGTAATATGAGCC 600
 Db 541 GACAAATCCCTTACCTTGAATGAGACCTACGAGCATCAGATGTTGTAATATGAGCC 600
 Qy 601 AAGCTCCAGCCCTGAGACTTCAAGAAAATGAGAGCAATCCAGAGCCGCTCAACAA 660
 Db 601 AAGCTCCAGCCCTGAGACTTCAAGAAAATGAGAGCAATCCAGAGCCGCTCAACAA 660
 Qy 661 TGGGTGCTCAATAGACCGAAGCCGGAATCACCAGATGTCATTCCTCCGAGCCATCAT 720
 Db 661 TGGGTGCTCAATAGACCGAAGCCGGAATCACCAGATGTCATTCCTCCGAGCCATCAT 720
 Qy 721 GAGCTCAGCTTCTGCTGCTGTTAAACCATTTTCAAGGCTGTTGAGAGTCAAG 780
 Db 721 GAGCTCAGCTTCTGCTGCTGTTAAACCATTTTCAAGGCTGTTGAGAGTCAAG 780
 Qy 781 TTCAGCCCTGAGAACACAGAGAAAGAACTGTTCTACAGGCTGATGAGAGTCTGTTCA 840
 Db 781 TTCAGCCCTGAGAACACAGAGAAAGAACTGTTCTACAGGCTGATGAGAGTCTGTTCA 840
 Qy 841 GCATCTATGATGTACCAAGAGGCAAGTTCGTTATCGCGCGCTGAGAGGCAACCCAG 900
 Db 841 GCATCTATGATGTACCAAGAGGCAAGTTCGTTATCGCGCGCTGAGAGGCAACCCAG 900
 Qy 901 GTCCTTGAATGCTGCTTCAAGAGTATGATCATCATAGTGTCTCATCTTGGCCCAAGCT 960
 Db 901 GTCCTTGAATGCTGCTTCAAGAGTATGATCATCATAGTGTCTCATCTTGGCCCAAGCT 960
 Qy 961 GAGAAAGACCTGCGCAAGGTGAGAGAAAGAACTCACCCAGAGTCTGAGAGTGGCTG 1020
 Db 961 GAGAAAGACCTGCGCAAGGTGAGAGAAAGAACTCACCCAGAGTCTGAGAGTGGCTG 1020
 Qy 1021 GATGAATGAGAGAGATGCTGCTGCTGTCATCATGCCCCGCTTCCGATGAGAGCGC 1080
 Db 1021 GATGAATGAGAGAGATGCTGCTGCTGTCATCATGCCCCGCTTCCGATGAGAGCGC 1080
 Qy 1081 TTCAGTTTGAAGAGACAGCTGCAAGCATAGGCTTTCGATCTGTTACCCCTGAAGA 1140
 Db 1081 TTCAGTTTGAAGAGACAGCTGCAAGCATAGGCTTTCGATCTGTTACCCCTGAAGA 1140
 Qy 1141 TCCAAACTCCAGATTTGTCAGAGAGCCGAGATGACCTTATGTCACAGATCATTC 1200
 Db 1141 TCCAAACTCCAGATTTGTCAGAGAGCCGAGATGACCTTATGTCACAGATCATTC 1200
 Qy 1201 CATAAAGCATTTCTTGAAGTAATAGAGAGGCAAGTGAAGCAGTGCAGATCCGCTGT 1260
 Db 1201 CATAAAGCATTTCTTGAAGTAATAGAGAGGCAAGTGAAGCAGTGCAGATCCGCTGT 1260
 Qy 1261 GTGATTTGCTGCGCTTCCCTAAACCCCAACAGGGTGACTTTCAAGGCCCAAGGCTTTC 1320
 Db 1261 GTGATTTGCTGCGCTTCCCTAAACCCCAACAGGGTGACTTTCAAGGCCCAAGGCTTTC 1320
 Qy 1321 CTGGTTTTTATAGAGAGATTTCTCTGAACACTATTTATCTTCAAGGGCAGATAGCAAC 1380
 Db 1321 CTGGTTTTTATAGAGAGATTTCTCTGAACACTATTTATCTTCAAGGGCAGATAGCAAC 1380
 Qy 1381 CTTGTGTGTTAAG 1392
 Db 1381 CTTGTGTGTTAAG 1392

RESULT 9
 HUMANITIV 1467 bp mRNA linear PRI 07-FEB-1999
 LOCUS Human mRNA for antithrombin III variant, complete cds.
 DEFINITION D29832
 ACCESSION D29832
 VERSION AT-III; antithrombin III.
 KEYWORDS Homo sapiens (individual-isolate AT-III Kyoto) cDNA to mRNA, clone pK16c.
 SOURCE ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (sites) Tsuji, H., Takada, O., Nakagawa, M., Tanaka, S. and Hashimoto-Gotoh, T.
 Hereditary antithrombin III deficiency: identification of an arginine-406 to methionine point mutation near protease reactive site
 (in) Yoshida, T.O. and Wilson, J.M. (Eds.);
 MOLECULAR APPROACHES TO THE STUDY AND TREATMENT OF HUMAN DISEASES: 51-53;
 Elsevier Science (1992)
 2 (bases 1 to 1467)
 Hashimoto-Gotoh, T.
 JOURNAL Unpublished
 FEATURES
 source location/Qualifiers
 1. 1467
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 22..1419
 /note="Wild type AT-III has 'G' instead of 'C' at position 1337 nt. Also amino acid residue changes from Met to Arg at position 406 aa in wild type AT-III."
 /codon_start=1
 /product="antithrombin III (AT-III) variant"
 /protein_id="BAA06212.1"
 /db_xref="gi:576554"
 /translation="MYSVITGVSGKRYVLSLILGFWDVCTCHSPVDICTAKR
 RDIENPWCITRSPKATDEDSBOKIPEATNNRWELSKASRFATTYOHLADS
 KNDNDLFLSLPLSISTAFAMTKLACNDITLQDLVEFDTISKTSQIHPFAKLN
 CRLVRKANKSKIVSANLFGDKSLTFNEVTDJSELVYAKLQPLPFKEAEQSRRA
 INKVASNTBGRIDVJPESEAINELTVLVNTYVFKELMSKESPEKRELEFYKAD
 GESCSVMYQEGKRYRYRVAEGVOVELPKRGDITVYLILPPEKSLATVEKLETP
 EVLDPMDELEEMNLVYHMRPRFIEDGSLSEQLQDMLVLPSPESKLEGIYAEGR
 DDLVSDAFHFAKFLVNEBESBAALSTAVIAGSLNENRVTFANMFLVFIKEVPL
 NTLIFMGVNAIPCVK"
 BASE COUNT 381 a 375 c 364 g 347 t
 ORIGIN
 Query Match 98.6%; Score 1375.6; DB 9; Length 1467;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1991; Conservative 0; Mismatches 4; Indels 3; Gaps 1;
 Qy 1 ATGTATTCATGATGATGAGAACTGTAACCTCTGGAAGAAAGAGATTATCTTTGTC 60
 Db 22 ATGTATTCATGATGATGAGAACTGTAACCTCTGGAAGAAAGAGATTATCTTTGTC 81
 Qy 61 TTGCTGCTATTGGCTTCTGGAAGCTGCTGACCTGTCACGGAGCCCTGAGACATCTGC 120
 Db 82 TTGCTGCTATTGGCTTCTGGAAGCTGCTGACCTGTCACGGAGCCCTGAGACATCTGC 141
 Qy 121 ACAGCCAGCCCGGGAGACATTCCCATGAAATCCCATGTGCATTATACCGCTCCCGAGAA 180
 Db 142 ACAGCCAGCCCGGGAGACATTCCCATGAAATCCCATGTGCATTATACCGCTCCCGAGAA 201
 Qy 181 AAGGCACTGAGAGATGAGAGGCTCAGAACAGAAAGATCCCGAGGGCCAC---CAACGGGGT 237
 Db 202 AAGGCACTGAGAGATGAGAGGCTCAGAACAGAAAGATCCCGAGGGCCACCAACCGGGCT 261
 Qy 238 GTCTGGAACTGTCAGAGCCCAATTCGCGCTTGTGTAACACTTTCTATCAGACCTGCA 297
 Db 262 GTCTGGAACTGTCAGAGCCCAATTCGCGCTTGTGTAACACTTTCTATCAGACCTGCA 321
 Qy 298 GATTCCAGAAATGACAAATGATTAATTTTCTGTGACCCCTGAGATATCTCCAGGCTTTT 357
 Db 322 GATTCCAGAAATGACAAATGATTAATTTTCTGTGACCCCTGAGATATCTCCAGGCTTTT 381
 Qy 358 GCTATGACCAAGCTGGGGGCTGTAATAGACACCCCTCCAGCAAGTATGAGAGATTTTAA 417
 Db 382 GCTATGACCAAGCTGGGGGCTGTAATAGACACCCCTCCAGCAAGTATGAGAGATTTTAA 441
 Qy 418 TTGACACCATATCTGAGAAACATCTGATGATGATCCACTTCTTTGCCAACTGAC 477
 Db 442 TTGACACCATATCTGAGAAACATCTGATGATGATCCACTTCTTTGCCAACTGAC 501


```

Ov  478 TGCAGCTCTATCGAAGCAAAATCTCCAGTAGTATACAGCCAAATGCCCTTTT 537
Db  502 TGCAGCTCTATCGAAGCAAAATCTCCAGTAGTATACAGCCAAATGCCCTTTT 561
Ov  538 GGAAGCAAAATCCCTTACCTTCAATGAGACTTACAGAGCATCAGTAGTGGTATATGA 597
Db  562 GGAAGCAAAATCCCTTACCTTCAATGAGACTTACAGAGCATCAGTAGTGGTATATGA 621
Ov  598 GCCAAGCTCGAGCCCTCGAGACTTCAAGAAAATCGAGCAATCCAGAGCCGATCAAC 657
Db  622 GCCAAGCTCGAGCCCTCGAGACTTCAAGAAAATCGAGCAATCCAGAGCCGATCAAC 681
Ov  658 AAATGGGTGTCATTAAGACCGAAGCCGATATCCAGATGTCTATCCCTCGAAGCCATC 717
Db  682 AAATGGGTGTCATTAAGACCGAAGCCGATATCCAGATGTCTATCCCTCGAAGCCATC 741
Ov  718 AATGAGCTCACTGTTCTGGTGTGTTTACCATTTACTTCAAGGGCTGTGGAAGTCA 777
Db  742 AATGAGCTCACTGTTCTGGTGTGTTTACCATTTACTTCAAGGGCTGTGGAAGTCA 801
Ov  778 AAGTTAGCCCTGAGAAACAAGAGAGAACTGTTCTACAAAGGCTGATGAGAGTGTGT 837
Db  802 AAGTTAGCCCTGAGAAACAAGAGAGAACTGTTCTACAAAGGCTGATGAGAGTGTGT 861
Ov  838 TCAGCATCTATGATGTACAGAGGCAAGTTCCGTTATCGCGCTGGAAGGCAAC 897
Db  862 TCAGCATCTATGATGTACAGAGGCAAGTTCCGTTATCGCGCTGGAAGGCAAC 921
Ov  898 CAGGTGCTTGAATGTCCTTCAAGAGTATGATATCAACATGGTCTCTATCTTGGCCAG 957
Db  922 CAGGTGCTTGAATGTCCTTCAAGAGTATGATATCAACATGGTCTCTATCTTGGCCAG 981
Ov  958 CCTGAGAAAGAGCTGCGCAAGGTGAGAGAACTACCCCAAGAGTGTGCGAGAGTGT 1017
Db  982 CCTGAGAAAGAGCTGCGCAAGGTGAGAGAACTACCCCAAGAGTGTGCGAGAGTGT 1041
Ov  1018 CTGAGATGAATGAGAGATGATGCTGTGTGTCACATGCCCCGCTTCCGATTTAGAGAC 1077
Db  1042 CTGAGATGAATGAGAGATGATGCTGTGTGTCACATGCCCCGCTTCCGATTTAGAGAC 1101
Ov  1078 GGTTCAGTTTGAAGAGAGAGTGCAGAGCATGGGCTTGTGATGCTGTACGCCCTGAA 1137
Db  1102 GGTTCAGTTTGAAGAGAGAGTGCAGAGCATGGGCTTGTGATGCTGTACGCCCTGAA 1161
Ov  1138 AAGTCCAAATCCCAAGTATGTTGAGAGAGCCGAGATGACTATGTCATGATGCA 1197
Db  1162 AAGTCCAAATCCCAAGTATGTTGAGAGAGCCGAGATGACTATGTCATGATGCA 1221
Ov  1198 TTTCATTAAGGCAATTTCTTGAAGTAAATGAAGAGGCAAGAGAGCTGCAAGTCCGT 1257
Db  1222 TTTCATTAAGGCAATTTCTTGAAGTAAATGAAGAGGCAAGAGAGCTGCAAGTCCGT 1281
Ov  1258 GTTGTGATTTGCGCGCTTGCATTAACCCCAAGGGGTGACTTTAAAGGCAACAGGCT 1317
Db  1282 GTTGTGATTTGCGCGCTTGCATTAACCCCAAGGGGTGACTTTAAAGGCAACAGGCT 1341
Ov  1318 TTTCCTGTTTTTAAAGAGAGTTCCTGTAACACTATTAATCTTCAAGGCAAGAGTACC 1377
Db  1342 TTTCCTGTTTTTAAAGAGAGTTCCTGTAACACTATTAATCTTCAAGGCAAGAGTACC 1401
Ov  1378 AACCTTGTGTTAAGTAA 1395
Db  1402 AACCTTGTGTTAAGTAA 1419

```

```

RESULT 10
OAAIII OAAIII 1751 bp mRNA linear MAM 19-JAN-1993
LOCUS O.aries mRNA for antithrombin III.
DEFINITION X68287
ACCESSION X68287.1 GI:1194
VERSION antithrombin III; coagulation inhibitor.
KEYWORDS

```

```

SOURCE
ORGANISM Ovis aries.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Caprine; Ovis.
REFERENCE
1 (bases 1 to 1751)
AUTHORS Niessen,R.W.L.M.
TITLE Direct Submission
JOURNAL Submitted (08-SEP-1992) R.W.L.M. Niessen, Center for Hemostasis,
Thrombosis, Atherosclerosis & Inflammation Research, Academic
Medical Center, Meibergdreef 9, 1105 AZ Amsterdam, THE NETHERLANDS
2 (bases 1 to 1751)
AUTHORS Niessen,R.W., Stuurk,A., Hordijk,P.L., Michiels,F. and Peters,M.
TITLE Sequence characterization of a sheep cDNA for antithrombin III
JOURNAL Biochim. Biophys. Acta 1171 (2), 207-210 (1992)
MEDLINE 93129622
PUBMED 1482684
FEATURES
source
location/Qualifiers
1..1751
/organism="Ovis aries"
/db_xref="taxon:9940"
/clone="PAT10, PAT15, PAT17"
/tissue_type="liver"
/clone_lib="sheep cDNA"
/dev_stage="adult"
51..1448
/codon_start=1
/product="antithrombin III"
/protein_id="CAA48347.1"
/db_xref="GI:1195"
/db_xref="SWISS-PROT:P32262"
/translation="MISGIGTIVTGRKSKMCLPFLLLIGLGCCTCHRSPYEDITAK
PRDIPVNMCIYRSPKKATEGSEBOKIPATNRVWELSKNSHFAFYQHLADS
KNNDNIDFLSLSTIAVAMTKLACNNNTLQMLEVFEDTISEKTSQIHFFPAKLN
CELYRKANSSELSVANSRLFGDKSITFNETYODISEVYGAALQPLDKNAGSRLT
INOMISNTEGRITDVIIPROAIDEFTVILVNTIYFKGLMKSPTSPEYTKXELFYKAD
GESCVPMYDEGKFRYRVAAGVOVLELPKKGDMITMPLILPKLEKLAVREPLT
DMLQWDELTEILLVMPHFRIEDSVEEQLODMLEDFSPKSRLLGCIYAEGR
NDLYVSDAFHKAFLFLEVEGSEAASTVISIAGSLNINRVTPQANRPFLVLRVAL
NTIIFMGVAPCVNA"
sig_peptide 51..146
mat_peptide 147..1445
BASE COUNT 467 a 422 c 431 g 431 t
ORIGIN
Query Match 79.9%; Score 1114.8; DB 4; Length 1751;
Best Local Similarity 87.8%; Pred. No. 1,1e-269;
Matches 1228; Conservative 0; Mismatches 167; Indels 3; Gaps 1;
Ov 1 ATGATTTCCATGTGATAGAGAACTGTAACTCTGGAAGAAAGAGAGCTTTATCTTTGTC 60
Db 51 ATGATTTCCATGTGATAGAGAACTGTAACTCTGGAAGAAAGAGAGCATGTGCTTTCCCT 110
Ov 61 TTGCTGCTCATTTGGCTTCTGGAAGCTGAGACCTGTCAAGGAGCCCTGT---GGACATC 117
Db 111 TTGCTGCTCATTTGGCTTCTGGAAGCTGAGACCTGTCAAGGAGCCCTGTGAGAGACATC 170
Ov 118 TGCACAGCAAGCCGCGGAGCAATTCCTCATGATCCCATGTGATTAACCGCTCCCGAG 177
Db 171 TGCACAGCAAGCCGCGGAGCAATTCCTGTGATCCCATGTGATTAACCGTTCCCAAG 230
Ov 178 AAGAGGCACTAGAGATGAGGGCTCAGAAAGAGATCCCGAGGCAACCAACGGGGT 237
Db 231 AAGAGGCACTAGAGGAGAGGGCTCAGAGAGAGATCCCTGGGGCCCAACACGGGGT 290
Ov 238 GTCTGGAACTGTCCAAAGGCAATTCGCGCTTGTATACACTTTATACAGACCTGGCA 297
Db 291 GTCTGGAACTGTCCAAAGGCAATTCGCGCTTGTATACAGATTTGGCA 350
Ov 298 GATTCGAAGATGACATGATTAATTTCTGTGACCCCTGAGATATCTCCAGGCTTTT 357
Db 351 GATTCGAAGATTAATGACATGATTTTCTGTGACCCCTGAGATATCTCCAGACTTTT 410

```


110 TG---GACATCTGCAGCAGCAGCGGGGACATTCGATGATTCATTCATTTACC 166
 65 TGGAGATATCTGCACTGCAAGGCCAGAGACATCCAGTATTCATTCATTTACC 124
 167 GCTCCCGGAGAGAGAGCACTGAGATGAGGGCTCAGAAC-----AGAGATCC 217
 125 GGAACCTTGAGAGAGAGAGAGCGGGGCTAGGGTCAAGAGAGAGAGAGATCC 184
 218 CGAGAGCCACCAACGGGGTGTCTGGGAATCTGCAAGGCCAATTCCTGTTGTTACCA 277
 185 CAGATCTACTATCTCCCGTGTCTGGGAGTTATCAAGGCCAACCCGCTTCGCAATGG 244
 278 CTCTTATCAGACCTGAGAGATTCAGAAATGACAAATGATTAACATTTCCGTACACC 337
 245 AATCTCAAAATACCTGGAGAGCTCAAGAAATGACAAATGAAATCTTCAATGACACC 304
 338 TGAGTATCTCCACGGCTTTTGTATGACCAAGCTGGTGCCTGTATGACACCTCCAGC 397
 305 TCAGATCTCCACAGCCTCTGGAGATGACCAACTGGTGGGTAAACACCTCCAGC 364
 398 AACTGATGAGATTTTAAGTTTGAACACATATCTGAGAAAACATCTGATCAGATCCACT 457
 365 AGCTCAGAGAGGTTTTCGGTTGACACCATTTCAAGAGAGACATCTGATCAGATCCACT 424
 458 TCTTTTGGCCAACTGAGCTGCGACTCTATCGAAAAGCCAAATCCTCCAAAGTTAG 517
 425 CTTTCTGCGCAAGCTGAAATGCGCGCTTTATAGAGAGCCAAATCTCAGAGCTGG 484
 518 TATCAGCCAAATCGCTTTTGGAGACAATCCCTTACCTTCAATGAGACTTACAGAGACA 577
 485 TGTGGGCAATGTCTCTTTGGGAGAAATCTCAGCTTCAATGAGACTTACAGAGACA 544
 578 TCAGTGAATTGATATGAGAGCCAAAGCTCCAGCCCTGAGCTTCAAGGAAAATGAGAGC 637
 545 TCAGTGAAGTGTATTCAGAGCCAAATCTGCGCTTCAAGACTTCAAGAAAAGCCGAGC 604
 638 AATCCAGAGCGGCCCATCAACAAATGGGTGTCATTAAGCCGAGAGCGGAATCAACGAT 697
 605 AGTCCAGAGCTCTAATTAACAGATGATGCTCAATTAAGCCGAGAGCGGAATCAACGAT 664
 698 TCATTCCTCGGAGAGCCATCAATGAGCTCACTGTTCTGGTGTGTTAACAACCATTTACT 757
 665 TGATCCCTGAGAGGAGCATTTGATGACCTCAGCTCTGGTGTGTTAACAACCATTTACT 724
 758 TCAGAGGCTGTGAGATCAAAATGTTCAAGCCTGAGAGACAAAGAGAACTGTTCTACA 817
 725 TTAAGGGCCACTGGAATCGAGTTCCCTGCTCAGAAACAAAGAGATTAAGTTCTATA 784
 818 AGGCTGATGAGAGATCGTGTCAAGATCTATGATTTACCAAGAGCAAGTTCCGTTATC 877
 785 GAGAGGATGAGAGATCTGTAGACACCCCATGATATTAACAGAGAAAGTTCCGTTATC 844
 878 GAGCGGTGGC---TGAAGGACCCAGAGTGTGTTGATGCTTCAAGGATGATGATCA 934
 845 TCTCATCCCAAGGACACAGCTGAGGTGTGAGACTACATACAAAGGAGATGATCA 904
 935 CCATGCTCTCATCTTGGCCAAAGCTGAGAGAGCCTGGCCAAAGTGAAGAGAACTCA 994
 905 CCATGCTCTCATCTTGGCCAAAGCTGAGAGAGCCTGGCCAAAGTGAAGAGAACTCA 964
 995 CCCGAGAGTGTGAGAGAGTGGCTGATGATTTGAGAGAGATGATGCTGCTGCTCA 1054
 965 CCTTGAAGAGCTGAGAGAGTGGCTGAGCTCTATGAGAGAGATCTGCTTCAATCA 1024
 1055 TGCCCGCTTCCGATTTGAGAGAGCGCTTCAAGTTTGAAGAGAGAGCTGCAAGCATGGCC 1114
 1025 TCCCGCGCTTCCGATTTGAGAGAGCAATTTCAAGAGAGAGAGAGCTGGGAGCATGGGCG 1084
 1115 TTGTGATCTGTTCAAGCCTGAAAAGTCAAACTCCAGGTATTTGTTGAGAGAGCGGAG 1174
 1085 TGGTGATCTCTTCAAGCCTGAAAAGTCAAACTCCAGGTATTTGTTGAGAGAGCGGAG 1144

1175 ATGACCTCTATGTTCTGAGATGATTCATTAAGGATTTCTGAGATTAATGAGAGGCA 1234
 1145 CAGACCTGTATGTCTGATGCTTTCCACAAAGGCTTCTTGAAGTGAATGAGAGGCA 1204
 1235 GTGAGAGAGCTGCAAGTACCGCTGTTGATGCTGAGCCGTTGCTTAAACCCCAAGGG 1294
 1205 GTGAGGCGGAGCTTCTACCGTATGATCTCGGGCGCTTCAATTCATGAAACAAA 1264
 1295 TGACTTCAAGGCCAACAGGCTTCTGTTTATTAAGAGAGTCTCTGTAACACTA 1354
 1265 TGATCTTCAATGCTTCAAGAGCGCTTCTGCTGCTCATTCGGAGAGTGGCATCAACGCA 1324
 1355 TTATCTCATGAGGAGAGTACCAACCTTG 1385
 1325 TCATCTCATGAGGCGCAATTCATTCCTG 1355

RESULT 15
 AF335570 1542 bp mRNA linear VRT 02-FEB-2002
 LOCUS Sphenodon punctatus antithrombin III (AtIII) mRNA, complete cds.
 DEFINITION
 ACCESSION AF335570
 VERSION AF335570.1 GI:18478578
 KEYWORDS
 SOURCE
 ORGANISM
 Sphenodon punctatus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Lepidodonta; Sphenodontia; Sphenodontidae; Sphenodon.
 REFERENCE
 1 (bases 1 to 1542)
 Metcalf, V.J., Brennan, S.O., George, P.M. and Chambers, G.K.
 The identification of liver-expressed genes from tuatara
 JOURNAL
 2 (bases 1 to 1542)
 Metcalf, V.J., Brennan, S.O., George, P.M. and Chambers, G.K.
 Direct Submission
 TITLE
 Sphenodon punctatus antithrombin III (AtIII) mRNA, complete cds.
 JOURNAL
 Submitted (11-JUN-2001) Department of Pathology, Christchurch
 School of Medicine, University of Otago, PO Box 4345, Christchurch,
 New Zealand

FEATURES
 source
 1..1542
 /location/Qualifiers
 /organism="Sphenodon punctatus"
 /db_xref="taxon:8508"
 1..1542
 /gene="AtIII"
 85..1443
 /gene="AtIII"
 /note="serine protease inhibitor"
 /codon_start=1
 /product="antithrombin III"
 /protein_id="AAU73207.1"
 /db_xref="GI:18478578"
 /translation="MHLPFLIVISIGVVAPOHHVPICTAKPRDIADVNPWCYRSPB
 KKLQEGEPQEKELPDFTNPRWELSKNSRPATVYKHLADSKDHGNIEMSPIS
 ISTAPMTKLGCSTLQOLMEVRFPTLSEKTSDOHFFPAKNCGLYKANKSSEL
 VSNARLFGKTLNFTNYONVSELVYAKCLPLQKPELSRLIYNMVAKTEKRI
 TDVLPBGITPEFTLVVNTIYFKGRKSEPTENTMKEKFRDEPILCMVPMYQEA
 RFRTIYPSNKVOYLEPYKDDITMVLIIIPSKTYLAEBVRDMSQLEKMSMOE
 VMISVLPKFTIEDSFSLKENLQWGLVDIFSPENALPGIITEDKDLVSDAFHRA
 FLEVNIEGSEPAASTAVMTGRSLSNMKVFNANKPFLVIREVAINSIIFMGVSDP
 CL"

BASE COUNT 408 a 394 c 376 g 364 t
 ORIGIN
 Query Match 49.5%; Score 690; DB 5; Length 1542;
 Best Local Similarity 69.6%; Pred. No. 9.6e-163;
 Matches 968; Conservative 0; Mismatches 410; Indels 12; Gaps 2;

18 AGGAAGTATACCTCTGGAAGAGGTTATCTTTGCTTGGCTGCTCATTTGCTT 77
 57 AAG 116
 78 CTGGAGCTGCTGACCTGTCAAGGAGCCCTGTGAGACATCTGACAGAGAGAGAGAG 137

Db 117 CTGGGAGTGTAGCTCCCGACGACCCCGTGATATCTGCATGCAAGCCAAAGGA 176
 QY 138 CATTCCTCATGATCCCATGATGATTTACGCTCCCGGAGAAAGGCAACTGAGATGA 197
 Db 177 CATGCTGTGAATCCCATGATGATCTATGCTAGCCCCGAGAAAGAACTGCAGAAAGGCA 236
 QY 198 GGGCTCAGAAC-----AGAAATCCCGAGGCGCCACCAACCGCGTGTCTGGAACT 248
 Db 237 GGGACCGGAGAGAGAAAGAGAGCTTCCAGACTTTACCAATCCCGGTCTGGAGAGCT 296
 QY 249 GTCCAGGCGCAATTCGCGCTTGTCTACACTTTCTATGACACTGCGAGATTTCCAGAA 308
 Db 297 CTCCAAAGGGAATTCGCTTGTGCAAGGTCTTCTATAGACACTAGCTGATTTCCAGAA 356
 QY 309 TGACAAATGATTAACATTTCTCTGACCCCTGAGTATCTCCACGCTTTTGTATGACAA 368
 Db 357 TGACCATGGCAATATCTTCAATGCGCCCTCAGCATCTCCACGCTTTTGCCATGACCA 416
 QY 369 GCTGGGTCTCTGATGACACCTCCAGCACTGATGAGGATTTTAAGTTTGAACCAT 428
 Db 417 GCTCGGGGCTGTGGAGTACTCTCCAGCAGCTCATGAGGTGTCCGGTTTGACACAT 476
 QY 429 ATCTGAGAAACATCTGATGATGATCACTTCTTTCGCAACTGAATGCGGACTTA 488
 Db 477 TTGGAGAAAGACTTGTGATGATGATCACTTCTTTCGCAAGTGAATGCGGCTTTA 536
 QY 489 TCGAAAGCCCAACAAATCTCTCAAGTTAGTATGACCAATCGCTTTTGTGAGCAAAATC 548
 Db 537 CAAGAAAGCCCAATTAATCTCTGAGGTGTGTGAGCAATCGCTCTTTGGGAGAAAGAC 596
 QY 549 CCTTACCTTCAATGACACTTACAGAGCATGATGAGTGTGTATGAGAACCAAGTCCA 608
 Db 597 CCTGAACCTTCAACGAGACCTTACAGATGTCAGTGAACATGATTATGAGCCCAACTTG 656
 QY 609 GCCCTCGAAGCTTCAAGAAATGACAGCAATCCAGAGGCGCATCAACAAATGAGTGT 668
 Db 657 CCTTTGAACCTTCAAGAAACCAAGATTTGTCCAGGACTTAATTTAACAATGAGTGT 716
 QY 669 CATTAAGCCGAGAGCGGATCAACGATGATTTCCCTCGAAGCCCATCATGAGCTCAC 728
 Db 717 AATTAAGACAGAAACGATTTACAGTGTGATCCAGAGGAGGATCAACGAGTTTAC 776
 QY 729 TGTCTGTGTCTGTGTTAACCACTTATCTTCAAGGCGCTGTGAGTCAAGTTCAGCCC 788
 Db 777 TGTCTTGTGTCTGTGTTAACCACTTATCTTAAAGGCGAGTGTGAGTGTCTTAC 836
 QY 789 TGAGAACACAAAGAGAACTGTCTTACAAAGCTGATGAGAGTGTGCTTCAAGCTTAT 848
 Db 837 CGAGAACACCAAGATGAGAAAGTTCTTCAAGATGAGCCCAATTTGTGATGTGCCAT 896
 QY 849 GATGTACCAAGAGGCAAGTTCCGTTATCGGCGCTGTGAAGCA--CCCAAGTGTCT 905
 Db 897 GATGTACCAAGAGGCGAGTTCCGTTATCTTCCCTAGCAACAGGTGTGAGTGTCT 956
 QY 906 TGAAGTGTCTTCAAGAGTGTGATGATCAACATGATGATGATGATGATGATGATGAT 965
 Db 957 GGAAGTGTCTTCAAGAGGATGATCAACATGATGATGATGATGATGATGATGATGAT 1016
 QY 966 GAGCTGTGCGCAAGTGTGAGAAAGTCAACCCAGAGAGTGTGAGAGTGTGATGA 1025
 Db 1017 AGTCTGTGCGGAGTGTGAGAGGAACTGATGTGAGAGAGTGTGAGAGTGTGAGCTC 1076
 QY 1026 ATTGAGAGAGATGATGTGTGTGATGATGATGATGATGATGATGATGATGATGAT 1085
 Db 1077 CATGCAAGAGATGATGTGTGTGATGATGATGATGATGATGATGATGATGATGAT 1136
 QY 1086 TTTGAGAGAGAGTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1145
 Db 1137 CCGAAGAGAGAGTGTGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1196
 QY 1146 ACTCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1205
 Db 1197 GCTGCCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1256

QY 1206 GGCATTTCTTGAAGTAAATGAAGAGGAGTGAAGCACTGCAAGTACCGCTTGTGAT 1265
 Db 1257 AGCTTCTTGAAGTGAATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1316
 QY 1266 TGTGCGCGTGTGCTTAAACCCCAAGAGTGTGATGATGATGATGATGATGATGATGAT 1325
 Db 1317 CACAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1376
 QY 1326 TTTTAAAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1385
 Db 1377 GCTCATGAGGAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1436
 QY 1386 TGTTAAGTAA 1395
 Db 1437 CCTTAATGA 1446

Search completed: June 26, 2003, 23:33:58
 Job time : 3604 secs